

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>A61K 48/00, C07H 21/04, C07K 14/00, 16/00, C12N 15/63, 15/85, 15/86</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/66178</b> <b>(43) International Publication Date:</b> 9 November 2000 (09.11.00)
<b>(21) International Application Number:</b> PCT/US00/11664 <b>(22) International Filing Date:</b> 28 April 2000 (28.04.00) <b>(30) Priority Data:</b> 60/132,021 30 April 1999 (30.04.99) US <b>(71) Applicant (for all designated States except US):</b> EMORY UNIVERSITY [US/US]; 1380 South Oxford Road, Atlanta, GA 30322 (US). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> BODEN, Scott, D. [US/US]; 2842 Cravey Drive, N.E., Atlanta, GA 30345 (US). HAIR, Gregory, A. [US/US]; 1887 Chrysler Drive, N.E., Atlanta, GA 30345 (US). <b>(74) Agents:</b> GARRETT, Arthur, S. et al.; Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P., 1300 I Street, N.W., Washington, DC 20005-3315 (US).		<b>(81) Designated States:</b> AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GI, GM, KE, LS, MW, SD, SI, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>
<b>(54) Title:</b> LIM MINERALIZATION PROTEIN SPLICE VARIANTS		
<b>(57) Abstract</b>  <p>The present invention is directed to isolated nucleic acid molecules that encode LIM mineralization protein, or LMP. The invention further provides vectors comprising splice variants of nucleotide sequences that encode LMP, as well as host cells comprising those vectors. Moreover, the present invention relates to methods of inducing bone formation by transfecting osteogenic precursor cells with an isolated nucleic acid molecule comprising a nucleotide sequence encoding splice variants of LIM mineralization protein. The transfection may occur <i>ex vivo</i> or <i>in vivo</i> by direct injection of virus or naked plasmid DNA. In a particular embodiment, the invention provides a method of fusing a spine by transfecting osteogenic precursor cells with an isolated nucleic acid molecule having a nucleotide sequence encoding LIM mineralization protein, admixing the transfected osteogenic precursor cells with a matrix and contacting the matrix with the spine. Finally, the invention relates to methods for inducing systemic bone formation by stable transfection of host cells with the vectors of the invention.</p>		

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## LIM Mineralization Protein Splice Variants

## BACKGROUND OF THE INVENTION

## 5 1. Field of the Invention

The field of the invention relates generally to osteogenic cells and the formation of bone and boney tissue in mammalian species. Specifically, the invention concerns a novel family of proteins, and nucleic acids encoding those proteins, that enhances the efficacy of bone mineralization *in vitro* and *in vivo*.

10 The invention provides methods for treating a variety of pathological conditions associated with bone and boney tissue, such as, for example, spine fusion, fracture repair and osteoporosis.

## 2. Description of the Related Art

15 Osteoblasts are thought to differentiate from pluripotent mesenchymal stem cells. The maturation of an osteoblast results in the secretion of an extracellular matrix which can mineralize and form bone. The regulation of this complex process is not well understood but is thought to involve a group of signaling glycoproteins known as bone morphogenetic proteins (BMPs). These  
20 proteins have been shown to be involved with embryonic dorsal-ventral patterning, limb bud development, and fracture repair in adult animals. B. L. Hogan, Genes & Develop., 10:1580 (1996). This group of transforming growth factor-beta superfamily secreted proteins has a spectrum of activities in a variety of cell types at different stages of differentiation; differences in  
25 physiological activity between these closely related molecules have not been clarified. D. M. Kingsley, Trends Genet., 10:16 (1994).

To better discern the unique physiological role of different BMP signaling proteins, we recently compared the potency of BMP-6 with that of BMP-2 and BMP-4, for inducing rat calvarial osteoblast differentiation. Boden *et al.*,  
30 Endocrinology, 137:3401 (1996). We studied this process in first passage (secondary) cultures of fetal rat calvaria that require BMP or glucocorticoid for

initiation of differentiation. In this model of membranous bone formation, glucocorticoid (GC) or a BMP will initiate differentiation to mineralized bone nodules capable of secreting osteocalcin, the osteoblast-specific protein. This secondary culture system is distinct from primary rat osteoblast cultures which  
5 undergo spontaneous differentiation. In this secondary system, glucocorticoid resulted in a ten-fold induction of BMP-6 mRNA and protein expression which was responsible for the enhancement of osteoblast differentiation. Boden *et al.*, Endocrinology, 138:2920 (1997).

In addition to extracellular signals, such as the BMPs, intracellular  
10 signals or regulatory molecules may also play a role in the cascade of events leading to formation of new bone. One broad class of intracellular regulatory molecules are the LIM proteins, which are so named because they possess a characteristic structural motif known as the LIM domain. The LIM domain is a cysteine-rich structural motif composed of two special zinc fingers that are  
15 joined by a 2-amino acid spacer. Some proteins have only LIM domains, while others contain a variety of additional functional domains. LIM proteins form a diverse group, which includes transcription factors and cytoskeletal proteins. The primary role of LIM domains appears to be in mediating protein-protein interactions, through the formation of dimers with identical or different LIM  
20 domains, or by binding distinct proteins.

In LIM homeodomain proteins, that is, proteins having both LIM domains and a homeodomain sequence, the LIM domains function as negative regulatory elements. LIM homeodomain proteins are involved in the control of cell lineage determination and the regulation of differentiation, although LIM-  
25 only proteins may have similar roles. LIM-only proteins are also implicated in the control of cell proliferation since several genes encoding such proteins are associated with oncogenic chromosome translocations.

Humans and other mammalian species are prone to diseases or injuries that require the processes of bone repair and/or regeneration. For example,  
30 treatment of fractures would be improved by new treatment regimens that could stimulate the natural bone repair mechanisms, thereby reducing the time



required for the fractured bone to heal. In another example, individuals afflicted with systemic bone disorders, such as osteoporosis, would benefit from treatment regimens that would results in systemic formation of new bone. Such treatment regimens would reduce the incidence of fractures arising from the loss of bone mass that is a characteristic of this disease.

For at least these reasons, extracellular factors, such as the BMPs, have been investigated for the purpose of using them to stimulate formation of new bone *in vivo*. Despite the early successes achieved with BMPs and other extracellular signalling molecules, their use entails a number of disadvantages. For example, relatively large doses of purified BMPs are required to enhance the production of new bone, thereby increasing the expense of such treatment methods. Furthermore, extracellular proteins are susceptible to degradation following their introduction into a host animal. In addition, because they are typically immunogenic, the possibility of stimulating an immune response to the administered proteins is ever present.

Due to such concerns, it would be desirable to have available treatment regimens that use an intracellular signalling molecule to induce new bone formation. Advances in the field of gene therapy now make it possible to introduce into osteogenic precursor cells, that is, cells involved in bone formation, or peripheral blood leukocytes, nucleotide fragments encoding intracellular signals that form part of the bone formation process. Gene therapy for bone formation offers a number of potential advantages: (1) lower production costs; (2) greater efficacy, compared to extracellular treatment regimens, due to the ability to achieve prolonged expression of the intracellular signal; (3) it would by-pass the possibility that treatment with extracellular signals might be hampered due to the presence of limiting numbers of receptors for those signals; (4) it permits the delivery of transfected potential osteoprogenitor cells directly to the site where localized bone formation is required; and (5) it would permit systemic bone formation, thereby providing a treatment regimen for osteoporosis and other metabolic bone diseases.

## SUMMARY OF THE INVENTION

The present invention seeks to overcome the drawbacks in the prior art  
5 by providing novel compositions and methods for inducing bone formation  
using an intracellular signalling molecule that participates early in the cascade  
of events that leads to bone formation. Applicants have discovered 10-  
4/RLMP (SEQ ID NO: 1, SEQ ID NO: 2), a novel LIM gene with a sequence  
originally isolated from stimulated rat calvarial osteoblast cultures. The gene  
10 has been cloned, sequenced and assayed for its ability to enhance the efficacy  
of bone mineralization *in vitro*. The protein RLMP affects mineralization of  
bone matrix as well as differentiation of cells into the osteoblast lineage. Unlike  
other known cytokines, for example, BMPs, RLMP is not a secreted protein,  
but is instead an intracellular signaling molecule. This feature has the  
15 advantage of providing intracellular signaling amplification as well as easier  
assessment of transfected cells. It is also suitable for more efficient and  
specific *in vivo* applications. Suitable clinical applications include enhancement  
of bone repair in fractures, bone defects, bone grafting, and normal  
homeostasis in patients presenting with osteoporosis.

20 Applicants have also cloned, sequenced and deduced the amino acid  
sequence of a corresponding human protein, named human LMP-1. The  
human protein demonstrates enhanced efficacy of bone mineralization *in vitro*  
and *in vivo*.

In addition, the applicants have characterized a truncated (short) version  
25 of LMP-1, termed HLMP-1s. This short version resulted from a point mutation  
in one source of a cDNA clone, providing a stop codon which truncated the  
protein. The short version (LMP-1s) is fully functional when expressed in cell  
culture and *in vivo*.

Using PCR analysis of human heart cDNA library, Applicants have  
30 identified two alternative splice variants (referred to as HLMP-2 and HLMP-3)  
that differ from HLMP-1 in a region between base pairs 325 and 444 in the

nucleotide sequence encoding HLMP-1. The HLMP-2 sequence has a 119 base pair deletion and an insertion of 17 base pairs in this region. Compared to HLMP-1, the nucleotide sequence encoding HLMP-3 has no deletions, but it does have the same 17 base pairs as HLMP-2, which are inserted at position  
5 444 in the HLMP-1 sequence.

Additional features and advantages of the invention will be set forth in the description which follows, and in part will be apparent from the description, or may be learned by practice of the invention. The objectives and other advantages of the invention will be realized and attained by the methods and  
10 compositions of matter particularly pointed out in the written description and claims hereof.

In one broad aspect, the invention relates to an isolated nucleic acid molecule comprising a nucleic acid sequence encoding any LIM mineralization protein, wherein the nucleic acid molecule hybridizes under standard  
15 conditions to a nucleic acid molecule complementary to the full length of SEQ. ID NO: 25, and wherein the molecule hybridizes under highly stringent conditions to a nucleic acid molecule complementary to the full length of SEQ. ID NO: 26. In a specific aspect, the isolated nucleic acid molecule encodes HLMP-1, HLMP-1s, RLMP, HLMP-2, or HLMP-3. In addition, the invention is  
20 directed to vectors comprising these nucleic acid molecules, as well as host cells comprising the vectors. In another specific aspect, the invention relates to the proteins themselves.

In a second broad aspect, the invention relates to antibody that is specific for LIM mineralization protein, including HLMP-1, HLMP-1s, RLMP,  
25 HLMP-2, and HLMP-3. In one specific aspect, the antibody is a polyclonal antibody. In another specific aspect, the antibody is a monoclonal antibody.

In a third broad aspect, the invention relates to method of inducing bone formation wherein osteogenic precursor cells are transfected with an isolated nucleic acid molecule comprising a nucleotide sequence encoding LIM  
30 mineralization protein. In one specific aspect, the isolated nucleic acid molecule is in a vector, which may be a plasmid or a virus, such as adenovirus

or retrovirus. The transfection may occur *ex vivo* or *in vivo* by direct injection of the isolated nucleic acid molecule. The transfected isolated nucleic acid molecule may encode HLMP-1, HLMP-1s, RLMP, HLMP-2, or HLMP-3.

In a further aspect, the invention relates to methods of fusing a spine by  
 5 transfecting osteogenic precursor cells with an isolated nucleic acid molecule having a nucleotide sequence encoding LIM mineralization protein, admixing the transfected osteogenic precursor cells with a matrix and contacting the matrix with the spine.

In yet another aspect, the invention relates to methods for inducing  
 10 systemic bone formation by stable transfection of host cells with the vectors of the invention.

It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are intended to provide further explanation of the invention as claimed.

15

#### ABBREVIATIONS AND DEFINITIONS

	BMP	Bone Morphogenetic Protein
	HLMP-1	Human LMP-1, also
20		designated as Human LIM
		Protein or HLMP
	HLMP-1s	Human LMP-1 Short
		(truncated) protein
	HLMPU	Human LIM Protein Unique
25		Region
	LMP	LIM mineralization protein
	MEM	Minimal essential medium
	Trm	Triamcinolone
	-GlyP	Beta-glycerolphosphate
30	RACE	Rapid Amplification of cDNA
		Ends

	RLMP	Rat LIM mineralization protein, also designated as RLMP-1
	RLMPU	Rat LIM Protein Unique Region
5	RNasin	RNase inhibitor
	ROB	Rat Osteoblast
	10-4	Clone containing cDNA sequence for RLMP (SEQ ID NO: 2)
10	UTR	Untranslated Region
	HLMP-2	Human LMP Splice Variant 2
	HLMP-3	Human LMP Splice Variant 3
15		

### DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel mammalian LIM proteins, herein  
 20 designated LIM mineralization proteins, or LMP. The invention relates more  
 particularly to human LMP, known as HLMP or HLMP-1, or alternative splice  
 variants of human LMP, which are known as HLMP-2 or HLMP-3. The  
 applicants have discovered that these proteins enhance bone mineralization in  
 mammalian cells grown *in vitro*. When produced in mammals, LMP also  
 25 induces bone formation *in vivo*.

*Ex vivo* transfection of bone marrow cells, osteogenic precursor cells,  
 peripheral blood leukocytes, or mesenchymal stem cells with nucleic acid that  
 encodes LMP or HLMP, followed by reimplantation of the transfected cells in  
 the donor, is suitable for treating a variety of bone-related disorders or injuries.  
 30 For example, one can use this method to: augment long bone fracture repair;  
 generate bone in segmental defects; provide a bone graft substitute for

fractures; facilitate tumor reconstruction or spine fusion; and provide a local treatment (by injection) for weak or osteoporotic bone, such as in osteoporosis of the hip, vertebrae, or wrist. Transfection with LMP or HLMP-encoding nucleic acid is also useful in: the percutaneous injection of transfected marrow  
5 cells to accelerate the repair of fractured long bones; treatment of delayed union or non-unions of long bone fractures or pseudoarthrosis of spine fusions; and for inducing new bone formation in avascular necrosis of the hip or knee.

In addition to *ex vivo*-based methods of gene therapy, transfection of a recombinant DNA vector comprising a nucleic acid sequence that encodes  
10 LMP or HLMP can be accomplished *in vivo*. When a DNA fragment that encodes LMP or HLMP is inserted into an appropriate viral vector, for example, an adenovirus vector, the viral construct can be injected directly into a body site where endochondral bone formation is desired. By using a direct, percutaneous injection to introduce the LMP or HLMP sequence stimulation of  
15 bone formation can be accomplished without the need for surgical intervention either to obtain bone marrow cells (to transfect *ex vivo*) or to reimplant them into the patient at the site where new bone is required. Alden *et al.*, Neurosurgical Focus (1998), have demonstrated the utility of a direct injection method of gene therapy using a cDNA that encodes BMP-2, which was cloned  
20 into an adenovirus vector.

It is also possible to carry out *in vivo* gene therapy by directly injecting into an appropriate body site, a naked, that is, unencapsulated, recombinant plasmid comprising a nucleic acid sequence that encodes HLMP. In this embodiment of the invention, transfection occurs when the naked plasmid DNA  
25 is taken up, or internalized, by the appropriate target cells, which have been described. As in the case of *in vivo* gene therapy using a viral construct, direct injection of naked plasmid DNA offers the advantage that little or no surgical intervention is required. Direct gene therapy, using naked plasmid DNA that encodes the endothelial cell mitogen VEGF (vascular endothelial growth  
30 factor), has been successfully demonstrated in human patients. Baumgartner *et al.*, Circulation, 97(12):1114-23 (1998).

By using an adenovirus vector to deliver LMP into osteogenic cells, transient expression of LMP is achieved. This occurs because adenovirus does not incorporate into the genome of target cells that are transfected. Transient expression of LMP, that is, expression that occurs during the lifetime  
5 of the transfected target cells, is sufficient to achieve the objects of the invention. Stable expression of LMP, however, can occur when a vector that incorporates into the genome of the target cell is used as a delivery vehicle. Retrovirus-based vectors, for example, are suitable for this purpose.

Stable expression of LMP is particularly useful for treating various  
10 systemic bone-related disorders, such as osteoporosis and osteogenesis imperfecta. For this embodiment of the invention, in addition to using a vector that integrates into the genome of the target cell to deliver an LMP-encoding nucleotide sequence into target cells, LMP expression is placed under the control of a regulatable promoter. For example, a promoter that is turned on by  
15 exposure to an exogenous inducing agent, such as tetracycline, is suitable. Using this approach, one can stimulate formation of new bone on a systemic basis by administering an effective amount of the exogenous inducing agent. Once a sufficient quantity of bone mass is achieved, administration of the exogenous inducing agent is discontinued. This process may be repeated as  
20 needed to replace bone mass lost, for example, as a consequence of osteoporosis.

Antibodies specific for HLMP are particularly suitable for use in methods for assaying the osteoinductive, that is, bone-forming, potential of patient cells. In this way one can identify patients at risk for slow or poor healing of bone  
25 repair. Also, HLMP-specific antibodies are suitable for use in marker assays to identify risk factors in bone degenerative diseases, such as, for example, osteoporosis.

Following well known and conventional methods, the genes of the present invention are prepared by ligation of nucleic acid segments that  
30 encode LMP to other nucleic acid sequences, such as cloning and/or expression vectors. Methods needed to construct and analyze these

recombinant vectors, for example, restriction endonuclease digests, cloning protocols, mutagenesis, organic synthesis of oligonucleotides and DNA sequencing, have been described. For DNA sequencing DNA, the dieoxyterminator method is the preferred.

5 Many treatises on recombinant DNA methods have been published, including Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, 2nd edition (1988), Davis *et al.*, Basic Methods in Molecular Biology, Elsevier (1986), and Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience (1988). These reference manuals are  
10 specifically incorporated by reference herein.

Primer-directed amplification of DNA or cDNA is a common step in the expression of the genes of this invention. It is typically performed by the polymerase chain reaction (PCR). PCR is described in U.S. Patent No. 4,800,159 to Mullis *et al.* and other published sources. The basic principle of  
15 PCR is the exponential replication of a DNA sequence by successive cycles of primer extension. The extension products of one primer, when hybridized to another primer, becomes a template for the synthesis of another nucleic acid molecule. The primer-template complexes act as substrate for DNA polymerase, which in performing its replication function, extends the primers.  
20 The conventional enzyme for PCR applications is the thermostable DNA polymerase isolated from *Thermus aquaticus*, or Taq DNA polymerase.

Numerous variations of the basic PCR method exist, and a particular procedure of choice in any given step needed to construct the recombinant vectors of this invention is readily performed by a skilled artisan. For example,  
25 to measure cellular expression of 10-4/RLMP, RNA is extracted and reverse transcribed under standard and well known procedures. The resulting cDNA is then analyzed for the appropriate mRNA sequence by PCR.

The gene encoding the LIM mineralization protein is expressed in an expression vector in a recombinant expression system. Of course, the  
30 constructed sequence need not be the same as the original, or its complimentary sequence, but instead may be any sequence determined by the



degeneracy of the DNA code that nonetheless expresses an LMP having bone forming activity. Conservative amino acid substitutions, or other modifications, such as the occurrence of an amino-terminal methionine residue, may also be employed.

- 5           A ribosome binding site active in the host expression system of choice is ligated to the 5' end of the chimeric LMP coding sequence, forming a synthetic gene. The synthetic gene can be inserted into any one of a large variety of vectors for expression by ligating to an appropriately linearized plasmid. A regulatable promoter, for example, the *E. coli* lac promoter, is also suitable for  
10 the expression of the chimeric coding sequences. Other suitable regulatable promoters include trp, tac, recA, T7 and lambda promoters.

DNA encoding LMP is transfected into recipient cells by one of several standard published procedures, for example, calcium phosphate precipitation, DEAE-Dextran, electroporation or protoplast fusion, to form stable  
15 transformants. Calcium phosphate precipitation is preferred, particularly when performed as follows.

DNAs are coprecipitated with calcium phosphate according to the method of Graham and Van Der, Virology, 52:456 (1973), before transfer into cells. An aliquot of 40-50 g of DNA, with salmon sperm or calf thymus DNA as  
20 a carrier, is used for  $0.5 \times 10^6$  cells plated on a 100 mm dish. The DNA is mixed with 0.5 ml of 2X Hepes solution (280 mM NaCl, 50 mM Hepes and 1.5 mM  $\text{Na}_2\text{HPO}_4$ , pH 7.0), to which an equal volume of 2x  $\text{CaCl}_2$  (250 mM  $\text{CaCl}_2$  and 10 mM Hepes, pH 7.0) is added. A white granular precipitate, appearing after 30-40 minutes, is evenly distributed dropwise on the cells, which are allowed to  
25 incubate for 4-16 hours at 37°C. The medium is removed and the cells shocked with 15% glycerol in PBS for 3 minutes. After removing the glycerol, the cells are fed with Dulbecco's Minimal Essential Medium (DMEM) containing 10% fetal bovine serum.

DNA can also be transfected using: the DEAE-Dextran methods of  
30 Kimura *et al.*, Virology, 49:394 (1972) and Sompayrac *et al.*, Proc. Natl. Acad. Sci. USA, 78:7575 (1981); the electroporation method of Potter, Proc. Natl.

Acad. Sci. USA, 81:7161 (1984); and the protoplast fusion method of Sandri-Goddin *et al.*, Molec. Cell. Biol., 1:743 (1981).

Phosphoramidite chemistry in solid phase is the preferred method for the organic synthesis of oligodeoxynucleotides and polydeoxynucleotides. In addition, many other organic synthesis methods are available. Those methods are readily adapted by those skilled in the art to the particular sequences of the invention.

The present invention also includes nucleic acid molecules that hybridize under standard conditions to any of the nucleic acid sequences encoding the LIM mineralization proteins of the invention. "Standard hybridization conditions" will vary with the size of the probe, the background and the concentration of the nucleic acid reagents, as well as the type of hybridization, for example, *in situ*, Southern blot, or hybridization of DNA-RNA hybrids (Northern blot). The determination of "standard hybridization conditions" is within the level of skill in the art. For example, see U.S. Patent 5,580,775 to Freneau *et al.*, herein incorporated by reference for this purpose. See also, Southern, E. M., J. Mol. Biol., 98:503 (1975), Alwine *et al.*, Meth. Enzymol., 68:220 (1979), and Sambrook *et al.*, Molecular Cloning: A laboratory Manual, 2nd edition, pp. 7.19-7.50, Cold Spring Harbor Press (1989).

One preferred set of standard hybridization conditions involves a blot that is prehybridized at 42°C for 2 hours in 50% formamide, 5X SSPE (150 mM NaCl, 10 mM Na H<sub>2</sub>PO<sub>4</sub> [pH 7.4], 1 mM EDTA [pH 8.0]), 5X Denhardt's solution (20 mg Ficoll, 20 mg polyvinylpyrrolidone and 20 mg BSA per 100 ml water), 10% dextran sulphate, 1% SDS and 100 g/ml salmon sperm DNA. A <sup>32</sup>P-labelled cDNA probe is added, and hybridization is continued for 14 hours. Afterward, the blot is washed twice with 2X SSPE, 0.1% SDS for 20 minutes at 22°C, followed by a 1 hour wash at 65°C in 0.1X SSPE, 0.1 %SDS. The blot is then dried and exposed to x-ray film for 5 days in the presence of an intensifying screen.

Under "highly stringent conditions," a probe will hybridize to its target sequence if those two sequences are substantially identical. As in the case of standard hybridization conditions, one of skill in the art can, given the level of skill in the art and the nature of the particular experiment, determine the  
5 conditions under which only substantially identical sequences will hybridize.

Another aspect of the invention includes the proteins encoded by the nucleic acid sequences. In still another embodiment, the invention relates to the identification of such proteins based on anti-LMP antibodies. In this embodiment, protein samples are prepared for Western blot analysis by lysing  
10 cells and separating the proteins by SDS-PAGE. The proteins are transferred to nitrocellulose by electroblotting as described by Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley and Sons (1987). After blocking the filter with instant nonfat dry milk (1 gm in 100 ml PBS), anti-LMP antibody is added to the filter and incubated for 1 hour at room temperature. The filter is  
15 washed thoroughly with phosphate buffered saline (PBS) and incubated with horseradish peroxidase (HRPO)-antibody conjugate for 1 hour at room temperature. The filter is again washed thoroughly with PBS and the antigen bands are identified by adding diaminobenzidine (DAB).

Monospecific antibodies are the reagent of choice in the present  
20 invention, and are specifically used to analyze patient cells for specific characteristics associated with the expression of LMP. "Monospecific antibody" as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for LMP. "Homogeneous binding" as used herein refers to the ability of the antibody  
25 species to bind to a specific antigen or epitope, such as those associated with LMP, as described above. Monospecific antibodies to LMP are purified from mammalian antisera containing antibodies reactive against LMP or are prepared as monoclonal antibodies reactive with LMP using the technique of Kohler and Milstein, Nature, 256:495-97 (1975). The LMP specific antibodies  
30 are raised by immunizing animals such as, for example, mice, rats, guinea

pigs, rabbits, goats or horses, with an appropriate concentration of LMP either with or without an immune adjuvant.

In this process, preimmune serum is collected prior to the first immunization. Each animal receives between about 0.1 mg and about 1000 mg of LMP associated with an acceptable immune adjuvant, if desired. Such acceptable adjuvants include, but are not limited to, Freund's complete, Freund's incomplete, alum-precipitate, water in oil emulsion containing *Corynebacterium parvum* and tRNA adjuvants. The initial immunization consists of LMP in, preferably, Freund's complete adjuvant injected at multiple sites either subcutaneously (SC), intraperitoneally (IP) or both. Each animal is bled at regular intervals, preferably weekly, to determine antibody titer. The animals may or may not receive booster injections following the initial immunization. Those animals receiving booster injections are generally given an equal amount of the antigen in Freund's incomplete adjuvant by the same route. Booster injections are given at about three week intervals until maximal titers are obtained. At about 7 days after each booster immunization or about weekly after a single immunization, the animals are bled, the serum collected, and aliquots are stored at about -20° C.

Monoclonal antibodies (mAb) reactive with LMP are prepared by immunizing inbred mice, preferably Balb/c mice, with LMP. The mice are immunized by the IP or SC route with about 0.1 mg to about 10 mg, preferably about 1 mg, of LMP in about 0.5 ml buffer or saline incorporated in an equal volume of an acceptable adjuvant, as discussed above. Freund's complete adjuvant is preferred. The mice receive an initial immunization on day 0 and are rested for about 3-30 weeks. Immunized mice are given one or more booster immunizations of about 0.1 to about 10 mg of LMP in a buffer solution such as phosphate buffered saline by the intravenous (IV) route. Lymphocytes from antibody-positive mice, preferably splenic lymphocytes, are obtained by removing the spleens from immunized mice by standard procedures known in the art. Hybridoma cells are produced by mixing the splenic lymphocytes with an appropriate fusion partner, preferably myeloma cells, under conditions

which will allow the formation of stable hybridomas. Fusion partners may include, but are not limited to: mouse myelomas P3/NS1/Ag 4-1; MPC-11; S-194 and Sp 2/0, with Sp 2/0 being preferred. The antibody producing cells and myeloma cells are fused in polyethylene glycol, about 1000 mol. wt., at  
5 concentrations from about 30% to about 50%. Fused hybridoma cells are selected by growth in hypoxanthine, thymidine and aminopterin in supplemented Dulbecco's Modified Eagles Medium (DMEM) by procedures known in the art. Supernatant fluids are collected from growth positive wells on about days 14, 18, and 21, and are screened for antibody production by an  
10 immunoassay such as solid phase immunoradioassay (SPIRA) using LMP as the antigen. The culture fluids are also tested in the Ouchterlony precipitation assay to determine the isotype of the mAb. Hybridoma cells from antibody positive wells are cloned by a technique such as the soft agar technique of MacPherson, "Soft Agar Techniques", in Tissue Culture Methods and  
15 Applications, Kruse and Paterson (eds.), Academic Press (1973). See, also, Harlow *et al.*, Antibodies: A Laboratory Manual, Cold Spring Laboratory (1988).

Monoclonal antibodies may also be produced *in vivo* by injection of pristane- primed Balb/c mice, approximately 0.5 ml per mouse, with about  
20  $2 \times 10^6$  to about  $6 \times 10^6$  hybridoma cells about 4 days after priming. Ascites fluid is collected at approximately 8-12 days after cell transfer and the monoclonal antibodies are purified by techniques known in the art.

*In vitro* production in anti-LMP mAb is carried out by growing the hybridoma cell line in DMEM containing about 2% fetal calf serum to obtain  
25 sufficient quantities of the specific mAb. The mAb are purified by techniques known in the art.

Antibody titers of ascites or hybridoma culture fluids are determined by various serological or immunological assays, which include, but are not limited to, precipitation, passive agglutination, enzyme-linked immunosorbent antibody  
30 (ELISA) technique and radioimmunoassay (RIA) techniques. Similar assays

are used to detect the presence of the LMP in body fluids or tissue and cell extracts.

It is readily apparent to those skilled in the art that the above described methods for producing monospecific antibodies may be utilized to produce  
5 antibodies specific for polypeptide fragments of LMP, full-length nascent LMP polypeptide, or variants or alleles thereof.

In another embodiment, the invention is directed to alternative splice variants of HLMP-1. PCR analysis of human heart cDNA revealed mRNA for two HLMP alternative splice variants, named HLMP-2 and HLMP-3, that differ  
10 from HLMP-1 in a region between base pairs 325 and 444 in the hLMP-1 sequence. The HLMP-2 sequence has a 119 base pair deletion and an insertion of 17 base pairs in this region. These changes preserve the reading frame, resulting in a 423 amino acid protein, which compared to HLMP-1, has a net loss of 34 amino acids (40 amino acids deleted plus 6 inserted amino  
15 acids). HLMP-2 contains the c-terminal LIM domains that are present in HLMP-1.

Compared to HLMP-1, HLMP-3 has no deletions, but it does have the same 17 base pair insertion at position 444. This insertion shifts the reading frame, causing a stop codon at base pairs 459-461. As a result, HLMP-3  
20 encodes a protein of 153 amino acids. This protein lacks the c-terminal LIM domains that are present in HLMP-1 and HLMP-2. The predicted size of the proteins encoded by HLMP-2 and HLMP-3 was confirmed by western blot analysis.

PCR analysis of the tissue distribution of the three splice variants  
25 revealed that they are differentially expressed, with specific isoforms predominating in different tissues. HLMP-1 is apparently the predominant form expressed in leukocytes, spleen, lung, placenta, and fetal liver. HLMP-2 appears to be the predominant isoform in skeletal muscle, bone marrow, and heart tissue. HLMP-3, however, was not the predominant isoform in any tissue  
30 examined.

Overexpression of HLMP-3 in secondary rat osteoblast cultures induced bone nodule formation ( $287 \pm 56$ ) similar to the effect seen for glucicorticoid ( $272 \pm 7$ ) and HLMP-1 ( $232 \pm 200$ ). Since HLMP-3 lacks the C-terminal LIM domains, these regions are not required for osteoinductive activity.

5 Overexpression of HLMP-2, however, did not induce nodule formation ( $11 \pm 3$ ). These data suggest that the amino acids encoded by the deleted 119 base pairs are necessary for osteoinduction. The data also suggest that the distribution of HLMP splice variants may be important for tissue-specific function. Surprisingly, we have shown that HLMP-2 inhibits steroid-induced

10 osteoblast formation in secondary rat osteoblast cultures. Therefore, HLMP-2 will have therapeutic utility in clinical situations where bone formation is not desirable.

On July 22, 1997, a sample of 10-4/RLMP in a vector designated pCMV2/RLMP (which is vector pRc/CMV2 with insert 10-4 clone/RLMP) was

15 deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD 20852. The culture accession number for that deposit is 209153. On March 19, 1998, a sample of the vector pHis-A with insert HLMP-1s was deposited at the American Type Culture Collection ("ATCC"). The culture accession number for that deposit is 209698. On April 14, 2000,

20 samples of plasmids pHAhLMP-2 (vector pHisA with cDNA insert derived from human heart muscle cDNA with HLMP-2) and pHAhLMP-3 (vector pHisA with cDNA insert derived from human heart muscle cDNA with HLMP-3) were deposited with the ATCC, 10801 University Blvd., Manassas, VA, 20110-2209, USA, under the conditions of the Budapest treaty. The accession numbers for

25 these deposits are \_\_\_\_\_ and \_\_\_\_\_, respectively. These deposits, as required by the Budapest Treaty, will be maintained in the ATCC for at least 30 years and will be made available to the public upon the grant of a patent disclosing them. It should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent

30 rights granted by government action.

In assessing the nucleic acids, proteins, or antibodies of the invention, enzyme assays, protein purification, and other conventional biochemical methods are employed. DNA and RNA are analyzed by Southern blotting and Northern blotting techniques, respectively. Typically, the samples analyzed are size fractionated by gel electrophoresis. The DNA or RNA in the gels are then transferred to nitrocellulose or nylon membranes. The blots, which are replicas of sample patterns in the gels, were then hybridized with probes. Typically, the probes are radiolabelled, preferably with  $^{32}\text{P}$ , although one could label the probes with other signal-generating molecules known to those in the art. Specific bands of interest can then be visualized by detection systems, such as autoradiography.

For purposes of illustrating preferred embodiments of the present invention, the following, non-limiting examples are included. These results demonstrate the feasibility of inducing or enhancing the formation of bone using the LIM mineralization proteins of the invention, and the isolated nucleic acid molecules encoding those proteins.

#### Example 1: Calvarial Cell Culture

Rat calvarial cells, also known as rat osteoblasts ("ROB"), were obtained from 20-day pre-parturition rats as previously described. Boden *et al.*, Endocrinology, 137(8):3401-07 (1996). Primary cultures were grown to confluence (7 days), trypsinized, and passed into 6-well plates ( $1 \times 10^5$  cells/35 mm well) as first subculture cells. The subculture cells, which were confluent at day 0, were grown for an additional 7 days. Beginning on day 0, media were changed and treatments (Tm and/or BMPs) were applied, under a laminar flow hood, every 3 or 4 days. The standard culture protocol was as follows: days 1-7, MEM, 10% FBS, 50 g/ml ascorbic acid,  $\pm$  stimulus; days 8-14, BGJb medium, 10% FBS, 5mM -GlyP (as a source of inorganic phosphate to permit mineralization). Endpoint analysis of bone nodule formation and osteocalcin secretion was performed at day 14. The dose of BMP was chosen as 50 ng/ml



based on pilot experiments in this system that demonstrated a mid-range effect on the dose-response curve for all BMPs studied.

#### EXAMPLE 2: Antisense Treatment and Cell Culture

5

To explore the potential functional role of LMP-1 during membranous bone formation, we synthesized an antisense oligonucleotide to block LMP-1 mRNA translation and treated secondary osteoblast cultures that were undergoing differentiation initiated by glucocorticoid. Inhibition of RLMP  
10 expression was accomplished with a highly specific antisense oligonucleotide (having no significant homologies to known rat sequences) corresponding to a 25 bp sequence spanning the putative translational start site (SEQ ID NO: 42). Control cultures either did not receive oligonucleotide or they received sense oligonucleotide. Experiments were performed in the presence (preincubation)  
15 and absence of lipofectamine. Briefly, 22 g of sense or antisense RLMP oligonucleotide was incubated in MEM for 45 minutes at room temperature. Following that incubation, either more MEM or pre-incubated lipofectamine/MEM (7% v/v; incubated 45 minutes at room temperature) was added to achieve an oligonucleotide concentration of 0.2 M. The resulting  
20 mixture was incubated for 15 minutes at room temperature. Oligonucleotide mixtures were then mixed with the appropriate medium, that is, MEM/Ascorbate/ $\pm$ Trm, to achieve a final oligonucleotide concentration of 0.1 M.

Cells were incubated with the appropriate medium ( $\pm$ stimulus) in the  
25 presence or absence of the appropriate oligonucleotides. Cultures originally incubated with lipofectamine were re-fed after 4 hours of incubation (37°C; 5% CO<sub>2</sub>) with media containing neither lipofectamine nor oligonucleotide. All cultures, especially cultures receiving oligonucleotide, were re-fed every 24 hours to maintain oligonucleotide levels.

30 LMP-1 antisense oligonucleotide inhibited mineralized nodule formation and osteocalcin secretion in a dose-dependent manner, similar to the effect of

BMP-6 oligonucleotide. The LMP-1 antisense block in osteoblast differentiation could not be rescued by addition of exogenous BMP-6, while the BMP-6 antisense oligonucleotide inhibition was reversed with addition of BMP-6. This experiment further confirmed the upstream position of LMP-1 relative to BMP-6 in the osteoblast differentiation pathway. LMP-1 antisense oligonucleotide also inhibited spontaneous osteoblast differentiation in primary rat osteoblast cultures.

### EXAMPLE 3: Quantitation of Mineralized Bone Nodule Formation

10

Cultures of ROB's prepared according to Examples 1 and 2 were fixed overnight in 70% ethanol and stained with von Kossa silver stain. A semi-automated computerized video image analysis system was used to quantitate nodule count and nodule area in each well. Boden *et al.*, Endocrinology, 137(8):3401-07 (1996). These values were then divided to calculate the area per nodule values. This automated process was validated against a manual counting technique and demonstrated a correlation coefficient of 0.92 ( $p < 0.000001$ ). All data are expressed as the mean  $\pm$  standard error of the mean (S.E.M.) calculated from 5 or 6 wells at each condition. Each experiment was confirmed at least twice using cells from different calvarial preparations.

15  
20

### EXAMPLE 4: Quantitation of Osteocalcin Secretion

Osteocalcin levels in the culture media were measured using a competitive radioimmunoassay with a monospecific polyclonal antibody (Pab) raised in our laboratory against the C-terminal nonapeptide of rat osteocalcin as described in Nanes *et al.*, Endocrinology, 127:588 (1990). Briefly, 1 g of nonapeptide was iodinated with 1 mCi  $^{125}\text{I}$ -Na by the lactoperoxidase method. Tubes containing 200  $\mu$ l of assay buffer (0.02 M sodium phosphate, 1 mM EDTA, 0.001% thimerosal, 0.025% BSA) received media taken from cell cultures or osteocalcin standards (0 - 12,000 fmole) at 100  $\mu$ l/tube in assay

25  
30

buffer. The Pab (1:40,000; 100 l) was then added, followed by the iodinated peptide (12,000 cpm; 100 l). Samples tested for non-specific binding were prepared similarly but contained no antibody.

Bound and free PABs were separated by the addition of 700 l goat anti-rabbit IgG, followed by incubation for 18 hours at 4°C. After samples were centrifuged at 1200 rpm for 45 minutes, the supernatants were decanted and the precipitates counted in a gamma counter. Osteocalcin values were reported in fmole/100 l, which was then converted to pmole/ml medium (3-day production) by dividing those values by 100. Values were expressed as the mean  $\pm$  S.E.M. of triplicate determinations for 5-6 wells for each condition. Each experiment was confirmed at least two times using cells from different calvarial preparations.

#### EXAMPLE 5: Effect of Trm and RLMP on Mineralization *In Vitro*

15

There was little apparent effect of either the sense or antisense oligonucleotides on the overall production of bone nodules in the non-stimulated cell culture system. When ROB's were stimulated with Trm, however, the antisense oligonucleotide to RLMP inhibited mineralization of nodules by > 95%. The addition of exogenous BMP-6 to the oligonucleotide-treated cultures did not rescue the mineralization of RLMP-antisense-treated nodules.

Osteocalcin has long been synonymous with bone mineralization, and osteocalcin levels have been correlated with nodule production and mineralization. The RLMP-antisense oligonucleotide significantly decreases osteocalcin production, but the nodule count in antisense-treated cultures does not change significantly. In this case, the addition of exogenous BMP-6 only rescued the production of osteocalcin in RLMP-antisense-treated cultures by 10-15%. This suggests that the action of RLMP is downstream of, and more specific than, BMP-6.

EXAMPLE 6: Harvest and Purification of RNA

- Cellular RNA from duplicate wells of ROBs (prepared according to
- 5 Examples 1 and 2 in 6-well culture dishes) was harvested using 4M guanidine isothiocyanate (GIT) solution to yield statistical triplicates. Briefly, culture supernatant was aspirated from the wells, which were then overlaid with 0.6 ml of GIT solution per duplicate well harvest. After adding the GIT solution, the plates were swirled for 5-10 seconds (being as consistent as possible).
- 10 Samples were saved at -70°C for up to 7 days before further processing.

- RNA was purified by a slight modification of standard methods according to Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Ed., chapter 7.19, Cold Spring Harbor Press (1989). Briefly, thawed samples received 60 l 2.0 M sodium acetate (pH 4.0), 550 l phenol (water saturated)
- 15 and 150 l chloroform:isoamyl alcohol (49:1). After vortexing, the samples were centrifuged (10000 x g; 20 minutes; 4°C), the aqueous phase transferred to a fresh tube, 600 l isopropanol was added and the RNA precipitated overnight at -20°C.

- Following the overnight incubation, the samples were centrifuged
- 20 (10000 x g; 20 minutes) and the supernatant was aspirated gently. The pellets were resuspended in 400 l DEPC-treated water, extracted once with phenol:chloroform (1:1), extracted with chloroform:isoamyl alcohol (24:1) and precipitated overnight at -20°C after addition of 40 l sodium acetate (3.0 M; pH 5.2) and 1.0 ml absolute ethanol. To recover the cellular RNA, the samples
- 25 were centrifuged (10000 x g; 20 min), washed once with 70% ethanol, air dried for 5-10 minutes and resuspended in 20 l of DEPC-treated water. RNA concentrations were calculated from optical densities that were determined with a spectrophotometer.

EXAMPLE 7: Reverse Transcription-Polymerase Chain Reaction

Heated total RNA (5  $\mu$ g in 10.5  $\mu$ l total volume DEPC-H<sub>2</sub>O at 65°C for 5 minutes) was added to tubes containing 4  $\mu$ l 5X MMLV-RT buffer, 2  $\mu$ l dNTPs, 2  $\mu$ l dT17 primer (10 pmol/ml), 0.5  $\mu$ l RNAsin (40U/ml) and 1  $\mu$ l MMLV-RT (200 units/ $\mu$ l). The samples were incubated at 37°C for 1 hour, then at 95°C for 5 minutes to inactivate the MMLV-RT. The samples were diluted by addition of 80  $\mu$ l of water.

Reverse-transcribed samples (5  $\mu$ l) were subjected to polymerase-chain reaction using standard methodologies (50  $\mu$ l total volume). Briefly, samples were added to tubes containing water and appropriate amounts of PCR buffer, 25 mM MgCl<sub>2</sub>, dNTPs, forward and reverse primers for glyceraldehyde 3-phosphate dehydrogenase (GAP, a housekeeping gene) and/or BMP-6, <sup>32</sup>P-dCTP, and Taq polymerase. Unless otherwise noted, primers were standardized to run consistently at 22 cycles (94°C, 30"; 58°C, 30"; 72°C, 20").

EXAMPLE 8: Quantitation of RT-PCR Products by Polyacrylamide Gel Electrophoresis (PAGE) and PhosphorImager Analysis

RT-PCR products received 5  $\mu$ l/tube loading dye, were mixed, heated at 65°C for 10 min and centrifuged. Ten  $\mu$ l of each reaction was subjected to PAGE (12% polyacrylamide:bis; 15 V/well; constant current) under standard conditions. Gels were then incubated in gel preserving buffer (10% v/v glycerol, 7% v/v acetic acid, 40% v/v methanol, 43% deionized water) for 30 minutes, dried (80°C) *in vacuo* for 1-2 hours and developed with an electronically-enhanced phosphorescence imaging system for 6-24 hours. Visualized bands were analyzed. Counts per band were plotted graphically.

EXAMPLE 9: Differential Display PCR

RNA was extracted from cells stimulated with glucocorticoid (Trm, 1 nM). Heated, DNase-treated total RNA (5 g in 10.5 l total volume in DEPC-H<sub>2</sub>O at 65°C for 5 minutes) was reverse transcribed as described in Example 7, but H-T<sub>11</sub>M (SEQ ID. NO: 4) was used as the MMLV-RT primer. The resulting cDNAs were PCR-amplified as described above, but with various commercial primer sets (for example, H-T<sub>11</sub>G (SEQ ID NO: 4) and H-AP-10 (SEQ ID. NO: 5); GenHunter Corp, Nashville, TN). Radiolabelled PCR products were fractionated by gel electrophoresis on a DNA sequencing gel. After electrophoresis, the resulting gels were dried *in vacuo* and autoradiographs were exposed overnight. Bands representing differentially-expressed cDNAs were excised from the gel and reamplified by PCR using the method of Conner *et al.*, Proc. Natl. Acad. Sci. USA, 88:278 (1983). The products of PCR reamplification were cloned into the vector PCR-II (TA cloning kit; InVitrogen, Carlsbad, CA).

EXAMPLE 10: Screening of a UMR 106 Rat Osteosarcoma Cell cDNA Library

A UMR 106 library (2.5 x 10<sup>10</sup> pfu/ml) was plated at 5 x 10<sup>4</sup> pfu/ml onto agar plates (LB bottom agar) and the plates were incubated overnight at 37°C. Filter membranes were overlaid onto plates for two minutes. Once removed, the filters were denatured, rinsed, dried and UV cross-linked. The filters were then incubated in pre-hybridization buffer (2X PIPES [pH 6.5], 5% formamide, 1% SDS and 100 g/ml denatured salmon sperm DNA) for 2 h at 42°C. A 260 base-pair radiolabelled probe (SEQ ID NO: 3; <sup>32</sup>P labelled by random priming) was added to the entire hybridization mix/filters, followed by hybridization for 18 hours at 42°C. The membranes were washed once at room temperature (10 min, 1 x SSC, 0.1% SDS) and three times at 55°C (15 min, 0.1 x SSC, 0.1% SDS).

After they were washed, the membranes were analyzed by autoradiography as described above. Positive clones were plaque purified. The procedure was repeated with a second filter for four minutes to minimize spurious positives. Plaque-purified clones were rescued as lambda SK(-) phagemids. Cloned cDNAs were sequenced as described below.

#### EXAMPLE 11: Sequencing of Clones

Cloned cDNA inserts were sequenced by standard methods. Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience (1988). Briefly, appropriate concentrations of termination mixture, template and reaction mixture were subjected to an appropriate cycling protocol (95°C,30s; 68°C,30s; 72°C,60s; x 25). Stop mixture was added to terminate the sequencing reactions. After heating at 92°C for 3 minutes, the samples were loaded onto a denaturing 6% polyacrylamide sequencing gel (29:1 acrylamide:bis-acrylamide). Samples were electrophoresed for about 4 hours at 60 volts, constant current. After electrophoresis, the gels were dried *in vacuo* and autoradiographed.

The autoradiographs were analyzed manually. The resulting sequences were screened against the databases maintained by the National Center for Biotechnology Information (NIH, Bethesda, MD; <http://www.ncbi.nlm.nih.gov/>) using the BLASTn program set with default parameters. Based on the sequence data, new sequencing primers were prepared and the process was repeated until the entire gene had been sequenced. All sequences were confirmed a minimum of three times in both orientations.

Nucleotide and amino acid sequences were also analyzed using the PCGENE software package (version 16.0). Per cent homology values for nucleotide sequences were calculated by the program NALIGN, using the following parameters: weight of non-matching nucleotides, 10; weight of non-matching gaps, 10; maximum number of nucleotides considered, 50; and minimum number of nucleotides considered, 50.

For amino acid sequences, per cent homology values were calculated using PALIGN. A value of 10 was selected for both the open gap cost and the unit gap cost.

5 EXAMPLE 12: Cloning of RLMP cDNA

The differential display PCR amplification products described in Example 9 contained a major band of approximately 260 base pairs. This sequence was used to screen a rat osteosarcoma (UMR 106) cDNA library.

- 10 Positive clones were subjected to nested primer analysis to obtain the primer sequences necessary for amplifying the full length cDNA: (SEQ. ID NOs: 11, 12, 29, 30 and 31) One of those positive clones selected for further study was designated clone 10-4.

- Sequence analysis of the full-length cDNA in clone 10-4, determined by  
15 nested primer analysis, showed that clone 10-4 contained the original 260 base-pair fragment identified by differential display PCR. Clone 10-4 (1696 base pairs; SEQ ID NO: 2) contains an open reading frame of 1371 base pairs encoding a protein having 457 amino acids (SEQ ID NO: 1). The termination codon, TGA, occurs at nucleotides 1444-1446. The polyadenylation signal at  
20 nucleotides 1675-1680, and adjacent poly(A)<sup>+</sup> tail, was present in the 3' noncoding region. There were two potential N-glycosylation sites, Asn-Lys-Thr and Asn-Arg-Thr, at amino acid positions 113-116 and 257-259 in SEQ ID NO: 1, respectively. Two potential cAMP- and cGMP-dependent protein kinase phosphorylation sites, Ser and Thr, were found at amino acid positions 191  
25 and 349, respectively. There were five potential protein kinase C phosphorylation sites, Ser or Thr, at amino acid positions 3, 115, 166, 219, 442. One potential ATP/GTP binding site motif A (P-loop), Gly-Gly-Ser-Asn-Asn-Gly-Lys-Thr, was determined at amino acid positions 272-279.

- In addition, two highly conserved putative LIM domains were found at  
30 amino acid positions 341-391 and 400-451. The putative LIM domains in this newly identified rat cDNA clone showed considerable homology with the LIM



domains of other known LIM proteins. However, the overall homology with other rat LIM proteins was less than 25%. RLMP (also designated 10-4) has 78.5% amino acid homology to the human enigma protein (see U.S. Patent No. 5,504,192), but only 24.5% and 22.7% amino acid homology to its closest rat homologs, CLP-36 and RIT-18, respectively.

#### EXAMPLE 13: Northern Blot Analysis of RLMP Expression

Thirty g of total RNA from ROB's, prepared according to Examples 1 and 2, was size fractionated by formaldehyde gel electrophoresis in 1% agarose flatbed gels and osmotically transblotted to nylon membranes. The blot was probed with a 600 base pair EcoR1 fragment of full-length 10-4 cDNA labeled with <sup>32</sup>P-dCTP by random priming.

Northern blot analysis showed a 1.7 kb mRNA species that hybridized with the RLMP probe. RLMP mRNA was up-regulated approximately 3.7-fold in ROB's after 24 hours exposure to BMP-6. No up-regulation of RMLP expression was seen in BMP-2 or BMP-4-stimulated ROB's at 24 hours.

#### EXAMPLE 14: Statistical Methods

For each reported nodule/osteocalcin result, data from 5-6 wells from a representative experiment were used to calculate the mean  $\pm$  S.E.M. Graphs may be shown with data normalized to the maximum value for each parameter to allow simultaneous graphing of nodule counts, mineralized areas and osteocalcin.

For each reported RT-PCR, RNase protection assay or Western blot analysis, data from triplicate samples of representative experiments, were used to determine the mean  $\pm$  S.E.M. Graphs may be shown normalized to either day 0 or negative controls and expressed as fold-increase above control values.

Statistical significance was evaluated using a one-way analysis of variance with post-hoc multiple comparison corrections of Bonferroni as appropriate. D. V. Huntsberger, "The Analysis of Variance," in Elements of Statistical Variance, P. Billingsley (ed.), pp. 298-330, Allyn & Bacon Inc.,  
5 Boston, MA (1977) and Sigmapstat, Jandel Scientific, Corte Madera, CA. Alpha levels for significance were defined as  $p < 0.05$ .

EXAMPLE 15: Detection of Rat LIM Mineralization Protein by Western Blot Analysis

10

Polyclonal antibodies were prepared according to the methods of England *et al.*, Biochim.Biophys. Acta, 623:171 (1980) and Timmer *et al.*, J. Biol. Chem., 268:24863 (1993).

HeLa cells were transfected with pCMV2/RLMP. Protein was harvested  
15 from the transfected cells according to the method of Hair *et al.*, Leukemia Research, 20:1 (1996). Western Blot Analysis of native RLMP was performed as described by Towbin *et al.*, Proc. Natl. Acad. Sci. USA, 76:4350 (1979).

EXAMPLE 16: Synthesis of the Rat LMP-Unique (RLMPU) derived Human  
20 PCR product

Based on the sequence of the rat LMP-1 cDNA, forward and reverse PCR primers (SEQ ID NOs: 15 and 16) were synthesized and a unique 223 base-pair sequence was PCR amplified from the rat LMP-1 cDNA. A similar  
25 PCR product was isolated from human MG63 osteosarcoma cell cDNA with the same PCR primers.

RNA was harvested from MG63 osteosarcoma cells grown in T-75 flasks. Culture supernatant was removed by aspiration and the flasks were overlayed with 3.0 ml of GIT solution per duplicate, swirled for 5-10 seconds,  
30 and the resulting solution was transferred to 1.5 ml eppendorf tubes (5 tubes with 0.6 ml/tube). RNA was purified by a slight modification of standard

methods, for example, see Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, chapter 7, page 19, Cold Spring Harbor Laboratory Press (1989) and Boden *et al.*, Endocrinology, 138:2820-28 (1997). Briefly, the 0.6 ml samples received 60  $\mu$ l 2.0 M sodium acetate (pH 4.0), 550  $\mu$ l water saturated phenol and 5 150  $\mu$ l chloroform:isoamyl alcohol (49:1). After addition of those reagents, the samples were vortexed, centrifuged (10000 x g; 20 min; 4°C) and the aqueous phase transferred to a fresh tube. Isopropanol (600  $\mu$ l) was added and the RNA was precipitated overnight at -20°C. The samples were centrifuged (10000 x g; 20 minutes) and the supernatant was aspirated gently. The pellets were 10 resuspended in 400  $\mu$ l of DEPC-treated water, extracted once with phenol:chloroform (1:1), extracted with chloroform:isoamyl alcohol (24:1) and precipitated overnight at -20°C in 40  $\mu$ l sodium acetate (3.0 M; pH 5.2) and 1.0 ml absolute ethanol. After precipitation, the samples were centrifuged (10000 x g; 20 min), washed once with 70% ethanol, air dried for 5-10 minutes and 15 resuspended in 20  $\mu$ l of DEPC-treated water. RNA concentrations were derived from optical densities.

Total RNA (5  $\mu$ g in 10.5  $\mu$ l total volume in DEPC-H<sub>2</sub>O) was heated at 65°C for 5 minutes, and then added to tubes containing 4  $\mu$ l 5X MMLV-RT buffer, 2  $\mu$ l dNTPs, 2  $\mu$ l dT17 primer (10 pmol/ml), 0.5  $\mu$ l RNAsin (40 U/ml) and 1  $\mu$ l MMLV-RT 20 (200 units/ $\mu$ l). The reactions were incubated at 37°C for 1 hour. Afterward, the MMLV-RT was inactivated by heating at 95°C for 5 minutes. The samples were diluted by addition of 80  $\mu$ l water.

Transcribed samples (5  $\mu$ l) were subjected to polymerase-chain reaction using standard methodologies (50  $\mu$ l total volume). Boden *et al.*, Endocrinology, 25 138:2820-28 (1997); Ausubel *et al.*, "Quantitation of rare DNAs by the polymerase chain reaction", in Current Protocols in Molecular Biology, chapter 15.31-1, Wiley & Sons, Trenton, NJ (1990). Briefly, samples were added to tubes containing water and appropriate amounts of PCR buffer (25 mM MgCl<sub>2</sub>, dNTPs, forward and reverse primers (for RLMPU; SEQ ID NOs: 15 and 16), 30 <sup>32</sup>P-dCTP, and DNA polymerase. Primers were designed to run consistently at

22 cycles for radioactive band detection and 33 cycles for amplification of PCR product for use as a screening probe (94°C, 30 sec, 58°C, 30 sec; 72°C, 20 sec).

Sequencing of the agarose gel-purified MG63 osteosarcoma-derived  
5 PCR product gave a sequence more than 95% homologous to the RLMPU PCR product. That sequence is designated HLMP unique region (HLMPU; SEQ ID NO: 6).

EXAMPLE 17: Screening of reverse-transcriptase-derived MG63 cDNA

10

Screening was performed with PCR using specific primers (SEQ ID NOs: 16 and 17) as described in Example 7. A 717 base-pair MG63 PCR product was agarose gel purified and sequenced with the given primers (SEQ. ID NOs: 12, 15, 16, 17, 18, 27 and 28). Sequences were confirmed a  
15 minimum of two times in both directions. The MG63 sequences were aligned against each other and then against the full-length rat LMP cDNA sequence to obtain a partial human LMP cDNA sequence (SEQ ID NO: 7).

EXAMPLE 18: Screening of a Human Heart cDNA Library

20

Based on Northern blot experiments, it was determined that LMP-1 is expressed at different levels by several different tissues, including human heart muscle. A human heart cDNA library was therefore examined. The library was plated at  $5 \times 10^4$  pfu/ml onto agar plates (LB bottom agar) and plates were  
25 grown overnight at 37° C. Filter membranes were overlaid onto the plates for two minutes. Afterward, the filters denatured, rinsed, dried, UV cross-linked and incubated in pre-hybridization buffer (2X PIPES [pH 6.5]; 5% formamide, 1% SDS, 100 g/ml denatured salmon sperm DNA) for 2 h at 42°C. A radiolabelled, LMP-unique, 223 base-pair probe ( $^{32}\text{P}$ , random primer labelling;  
30 SEQ ID NO: 6) was added and hybridized for 18 h at 42°C. Following hybridization, the membranes were washed once at room temperature (10 min,

1 x SSC, 0.1% SDS) and three times at 55°C (15 min, 0.1 x SSC, 0.1% SDS). Double-positive plaque-purified heart library clones, identified by autoradiography, were rescued as lambda phagemids according to the manufacturers' protocols (Stratagene, La Jolla, CA).

5           Restriction digests of positive clones yielded cDNA inserts of varying sizes. Inserts greater than 600 base-pairs in length were selected for initial screening by sequencing. Those inserts were sequenced by standard methods as described in Example 11.

          One clone, number 7, was also subjected to automated sequence  
10   analysis using primers corresponding to SEQ ID NOs: 11-14, 16 and 27. The sequences obtained by these methods were routinely 97-100% homologous. Clone 7 (Partial Human LMP-1 cDNA from a heart library; SEQ. ID NO: 8) contained sequence that was more than 87% homologous to the rat LMP cDNA sequence in the translated region.

15

EXAMPLE 19: Determination of Full-Length Human LMP-1 cDNA

          Overlapping regions of the MG63 human osteosarcoma cell cDNA sequence and the human heart cDNA clone 7 sequence were used to align  
20   those two sequences and derive a complete human cDNA sequence of 1644 base-pairs. NALIGN, a program in the PCGENE software package, was used to align the two sequences. The overlapping regions of the two sequences constituted approximately 360 base-pairs having complete homology except for a single nucleotide substitution at nucleotide 672 in the MG63 cDNA (SEQ ID  
25   NO: 7) with clone 7 having an "A" instead of a "G" at the corresponding nucleotide 516 (SEQ ID NO: 8).

          The two aligned sequences were joined using SEQIN, another subprogram of PCGENE, using the "G" substitution of the MG63 osteosarcoma cDNA clone. The resulting sequence is shown in SEQ ID NO: 9. Alignment of  
30   the novel human-derived sequence with the rat LMP-1 cDNA was accomplished with NALIGN. The full-length human LMP-1 cDNA sequence

(SEQ. ID NO: 9) is 87.3% homologous to the translated portion of rat LMP-1 cDNA sequence.

EXAMPLE 20: Determination of Amino Acid Sequence of Human LMP-1

5

The putative amino acid sequence of human LMP-1 was determined with the PCGENE subprogram TRANSL. The open reading frame in SEQ ID NO: 9 encodes a protein comprising 457 amino acids (SEQ. ID NO: 10). Using the PCGENE subprogram Palign, the human LMP-1 amino acid sequence was found to be 94.1% homologous to the rat LMP-1 amino acid sequence.

EXAMPLE 21: Determination of the 5 Prime Untranslated Region of the Human LMP cDNA

15

MG63 5' cDNA was amplified by nested RT-PCR of MG63 total RNA using a 5' rapid amplification of cDNA ends (5' RACE) protocol. This method included first strand cDNA synthesis using a lock-docking oligo (dT) primer with two degenerate nucleotide positions at the 3' end (Chenchik *et al.*, CLONTECHniques, X:5 (1995); Borson *et al.*, PC Methods Applic., 2:144 (1993)). Second-strand synthesis is performed according to the method of Gubler *et al.*, Gene, 25:263 (1983), with a cocktail of *Escherichia coli* DNA polymerase I, RNase H, and *E. coli* DNA ligase. After creation of blunt ends with T4 DNA polymerase, double-stranded cDNA was ligated to the fragment (5' -CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCGGGCAGGT- 3') (SEQ.ID NO: 19). Prior to RACE, the adaptor-ligated cDNA was diluted to a concentration suitable for Marathon RACE reactions (1:50). Adaptor-ligated double-stranded cDNA was then ready to be specifically cloned.

First-round PCR was performed with the adaptor-specific oligonucleotide, 5'-CCATCCTAATACGACTCACTATAGGGC- 3' (AP1) (SEQ.ID NO: 20) as sense primer and a Gene Specific Primer (GSP) from the unique region described in Example 16 (HLMPU). The second round of PCR was

30

performed using a nested primers GSP1-HLMPU (antisense/reverse primer) (SEQ. ID NO: 23) and GSP2-HLMPUF (SEQ. ID NO: 24) (see Example 16; sense/forward primer). PCR was performed using a commercial kit (Advantage cDNA PCR core kit; CloneTech Laboratories Inc., Palo Alto, CA) that utilizes  
5 an antibody-mediated, but otherwise standard, hot-start protocol. PCR conditions for MG63 cDNA included an initial hot-start denaturation (94°C, 60 sec) followed by: 94°C, 30 sec; 60°C, 30 sec; 68°C, 4 min; 30 cycles. The first-round PCR product was approximately 750 base-pairs in length whereas the nested PCR product was approximately 230 base-pairs. The first-round PCR  
10 product was cloned into linearized pCR 2.1 vector (3.9 Kb). The inserts were sequenced in both directions using M13 Forward and Reverse primers (SEQ. ID NO: 11; SEQ. ID NO: 12)

EXAMPLE 22: Determination of Full-length Human LMP-1 cDNA with 5 Prime  
15 UTR

Overlapping MG63 human osteosarcoma cell cDNA 5'-UTR sequence (SEQ ID NO: 21), MG63 717 base-pair sequence (Example 17; SEQ ID NO: 8) and human heart cDNA clone 7 sequence (Example 18) were aligned to derive  
20 a novel human cDNA sequence of 1704 base-pairs (SEQ.ID NO: 22). The alignment was accomplished with NALIGN, (both PCGENE and Omega 1.0; Intelligenetics). Over-lapping sequences constituted nearly the entire 717 base-pair region (Example 17) with 100% homology. Joining of the aligned sequences was accomplished with SEQIN.

25

EXAMPLE 23: Construction of LIM Protein Expression Vector

The construction of pHIS-5ATG LMP-1s expression vector was carried out with the sequences described in Examples 17 and 18. The 717 base-pair  
30 clone (Example 17; SEQ ID NO: 7) was digested with ClaI and EcoRV. A small fragment (~250 base-pairs) was gel purified. Clone 7 (Example 18; SEQ

ID NO: 8) was digested with Clal and XbaI and a 1400 base-pair fragment was gel purified. The isolated 250 base-pair and 1400 base-pair restriction fragments were ligated to form a fragment of ~1650 base-pairs.

Due to the single nucleotide substitution in Clone 7 (relative to the 717  
5 base-pair PCR sequence and the original rat sequence) a stop codon at translated base-pair 672 resulted. Because of this stop codon, a truncated (short) protein was encoded, hence the name LMP-1s. This was the construct used in the expression vector (SEQ ID NO: 32). The full length cDNA  
10 NO: 32 with the 5' RACE sequence (SEQ ID NO: 21). The amino acid sequence of LMP-1s (SEQ ID NO: 34) was then deduced as a 223 amino acid protein and confirmed by Western blot (as in Example 15) to run at the predicted molecular weight of ~23.7 kD.

The pHIS-ATG vector (InVitrogen, Carlsbad, CA) was digested with  
15 EcoRV and XbaI. The vector was recovered and the 1650 base-pair restriction fragment was then ligated into the linearized pHIS-ATG. The ligated product was cloned and amplified. The pHIS-ATG-LMP-1s Expression vector, also designated pHIS-A with insert HLMP-1s, was purified by standard methods.

20 EXAMPLE 24: Induction of Bone Nodule Formation and Mineralization *In vitro* with LMP Expression Vector

Rat Calvarial cells were isolated and grown in secondary culture according to Example 1. Cultures were either unstimulated or stimulated with  
25 glucocorticoid (GC) as described in Example 1. A modification of the Superfect Reagent (Qiagen, Valencia, CA) transfection protocol was used to transfect 3 g/well of each vector into secondary rat calvarial osteoblast cultures according to Example 25.

Mineralized nodules were visualized by Von Kossa staining, as described in  
30 Example 3. Human LMP-1s gene product overexpression alone induced bone nodule formation (~203 nodules/well) *in vitro*. Levels of nodules were



approximately 50% of those induced by the GC positive control (~412 nodules/well). Other positive controls included the pHisA-LMP-Rat expression vector (~152 nodules/well) and the pCMV2/LMP-Rat-Fwd Expression vector (~206 nodules/well), whereas the negative controls included the pCMV2/LMP-Rat-Rev. Expression vector (~2 nodules/well) and untreated (NT) plates (~4 nodules/well). These data demonstrate that the human cDNA was at least as osteoinductive as the rat cDNA. The effect was less than that observed with GC stimulation, most likely due to suboptimal doses of Expression vector.

10

#### EXAMPLE 25: LMP-Induced Cell Differentiation *In Vitro* and *In Vivo*

The rat LMP cDNA in clone 10-4 (see Example 12) was excised from the vector by double-digesting the clone with NotI and ApaI overnight at 37°C. Vector pCMV2 MCS (InVitrogen, Carlsbad, CA) was digested with the same restriction enzymes. Both the linear cDNA fragment from clone 10-4 and pCMV2 were gel purified, extracted and ligated with T4 ligase. The ligated DNA was gel purified, extracted and used to transform *E. coli* JM109 cells for amplification. Positive agar colonies were picked, digested with NotI and ApaI and the restriction digests were examined by gel electrophoresis. Stock cultures were prepared of positive clones.

A reverse vector was prepared in analogous fashion except that the restriction enzymes used were XbaI and HindIII. Because these restriction enzymes were used, the LMP cDNA fragment from clone 10-4 was inserted into pRc/CMV2 in the reverse (that is, non-translatable) orientation. The recombinant vector produced is designated pCMV2/RLMP.

An appropriate volume of pCMV10-4 (60 nM final concentration is optimal [3 g]; for this experiment a range of 0-600 nM/well [0-30 g/well] final concentration is preferred) was resuspended in Minimal Eagle Media (MEM) to 450 l final volume and vortexed for 10 seconds. Superfect was added (7.5 l/ml final solution), the solution was vortexed for 10 seconds and then incubated at

room temperature for 10 minutes. Following this incubation, MEM supplemented with 10% FBS (1 ml/well; 6 ml/plate) was added and mixed by pipetting.

The resulting solution was then promptly pipetted (1 ml/well) onto washed ROB cultures. The cultures were incubated for 2 hours at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. Afterward, the cells were gently washed once with sterile PBS and the appropriate normal incubation medium was added.

Results demonstrated significant bone nodule formation in all rat cell cultures which were induced with pCMV10-4. For example, pCMV10-4 transfected cells produced 429 nodules/well. Positive control cultures, which were exposed to Trm, produced 460 nodules/well. In contrast, negative controls, which received no treatment, produced 1 nodule/well. Similarly, when cultures were transfected with pCMV10-4 (reverse), no nodules were observed.

For demonstrating *de novo* bone formation *in vivo*, marrow was aspirated from the hindlimbs of 4-5 week old normal rats (rnu/+; heterozygous for recessive athymic condition). The aspirated marrow cells were washed in alpha MEM, centrifuged, and RBCs were lysed by resuspending the pellet in 0.83% NH<sub>4</sub>Cl in 10 mM Tris (pH 7.4). The remaining marrow cells were washed 3x with MEM and transfected for 2 hours with 9 µg of pCMV-LMP-1s (forward or reverse orientation) per 3 x 10<sup>6</sup> cells. The transfected cells were then washed 2X with MEM and resuspended at a concentration of 3 x 10<sup>7</sup> cells/ml.

The cell suspension (100 µl) was applied via sterile pipette to a sterile 2 x 5 mm type I bovine collagen disc (Sulzer Orthopaedics, Wheat Ridge, CO). The discs were surgically implanted subcutaneously on the skull, chest, abdomen or dorsal spine of 4-5 week old athymic rats (rnu/rnu). The animals were scarified at 3-4 weeks, at which time the discs or surgical areas were excised and fixed in 70% ethanol. The fixed specimens were analyzed by radiography and undecalcified histologic examination was performed on 5 mm

thick sections stained with Goldner Trichrome. Experiments were also performed using devitalized (guanidine extracted) demineralized bone matrix (Osteotech, Shrewsbury, NJ) in place of collagen discs.

Radiography revealed a high level of mineralized bone formation that conformed to the form of the original collagen disc containing LMP-1s transfected marrow cells. No mineralized bone formation was observed in the negative control (cells transfected with a reverse-oriented version of the LMP-1s cDNA that did not code for a translated protein), and absorption of the carrier appeared to be well underway.

Histology revealed new bone trabeculae lined with osteoblasts in the LMP-1s transfected implants. No bone was seen along with partial resorption of the carrier in the negative controls.

Radiography of a further experiment in which 18 sets (9 negative control pCMV-LMP-REV & 9 experimental pCMV-LMP-1s) of implants were added to sites alternating between lumbar and thoracic spine in athymic rats demonstrated 0/9 negative control implants exhibiting bone formation (spine fusion) between vertebrae. All nine of the pCMV-LMP-1s treated implants exhibited solid bone fusions between vertebrae.

EXAMPLE 26: The Synthesis of pHIS-5' ATG LMP-1s Expression Vector from the sequences Demonstrated in Examples 2 and 3.

The 717 base-pair clone (Example 17) was digested with Clal and EcoRV (New England Biologicals, city, MA). A small fragment (~250 base-pairs) was gel purified. Clone No. 7 (Example 18) was digested with Clal and XbaI. A 1400 base-pair fragment was gel purified from that digest. The isolated 250 base-pair and 1400 base-pair cDNA fragments were ligated by standard methods to form a fragment of ~1650 bp. The pHIS-A vector (InVitrogen) was digested with EcoRV and XbaI. The linearized vector was recovered and ligated to the chimeric 1650 base-pair cDNA fragment. The ligated product was cloned and amplified by standard methods, and the pHIS-

A-5' ATG LMP-1s expression vector, also denominated as the vector pHis-A with insert HLMP-1s, was deposited at the ATCC as previously described.

EXAMPLE 27: The Induction of Bone Nodule Formation and Mineralization *In*

5 *Vitro* With pHis-5' ATG LMP-1s Expression Vector

Rat calvarial cells were isolated and grown in secondary culture according to Example 1. Cultures were either unstimulated or stimulated with glucocorticoid (GC) according to Example 1. The cultures were transfected  
10 with 3 g of recombinant pHis-A vector DNA/well as described in Example 25. Mineralized nodules were visualized by Von Kossa staining according to Example 3.

Human LMP-1s gene product overexpression alone (*i.e.*, without GC stimulation) induced significant bone nodule formation (~203 nodules/well) *in*  
15 *vitro*. This is approximately 50% of the amount of nodules produced by cells exposed to the GC positive control (~412 nodules/well). Similar results were obtained with cultures transfected with pHisA-LMP-Rat Expression vector (~152 nodules/well) and pCMV2/LMP-Rat-Fwd (~206 nodules/well). In contrast, the negative control pCMV2/LMP-Rat-Rev yielded (~2 nodules/well),  
20 while approximately 4 nodules/well were seen in the untreated plates. These data demonstrate that the human LMP-1 cDNA was at least as osteoinductive as the rat LMP-1 cDNA in this model system. The effect in this experiment was less than that observed with GC stimulation; but in some the effect was comparable.

25

EXAMPLE 28: LMP Induces Secretion of a Soluble Osteoinductive Factor

Overexpression of RLMP-1 or HLMP-1s in rat calvarial osteoblast cultures as described in Example 24 resulted in significantly greater nodule  
30 formation than was observed in the negative control. To study the mechanism of action of LIM mineralization protein conditioned medium was harvested at

different time points, concentrated to 10 X, sterile filtered, diluted to its original concentration in medium containing fresh serum, and applied for four days to untransfected cells.

Conditioned media harvested from cells transfected with RLMP-1 or  
5 HLMP-1s at day 4 was approximately as effective in inducing nodule formation as direct overexpression of RLMP-1 in transfected cells. Conditioned media from cells transfected with RLMP-1 or HLMP-1 in the reverse orientation had no apparent effect on nodule formation. Nor did conditioned media harvested from LMP-1 transfected cultures before day 4 induce nodule formation. These  
10 data suggest that expression of LMP-1 caused the synthesis and/or secretion of a soluble factor, which did not appear in culture medium in effective amounts until 4 days post transfection.

Since overexpression of rLMP-1 resulted in the secretion of an osteoinductive factor into the medium, Western blot analysis was used to  
15 determine if LMP-1 protein was present in the medium. The presence of rLMP-1 protein was assessed using antibody specific for LMP-1 (QDPDEE) and detected by conventional means. LMP-1 protein was found only in the cell layer of the culture and not detected in the medium.

Partial purification of the osteoinductive soluble factor was accomplished  
20 by standard 25% and 100% ammonium sulfate cuts followed by DE-52 anion exchange batch chromatography (100 mM or 500 mM NaCl). All activity was observed in the high ammonium sulfate, high NaCl fractions. Such localization is consistent with the possibility of a single factor being responsible for conditioning the medium.

25

#### EXAMPLE 29: Gene Therapy In Lumbar Spine Fusion Mediated by Low Dose Adenovirus

This study determined the optimal dose of adenoviral delivery of the  
30 LMP-1 cDNA (SEQ ID NO: 2) to promote spine fusion in normal, that is, immune competent, rabbits.

A replication-deficient human recombinant adenovirus was constructed with the LMP-1 cDNA (SEQ ID NO: 2) driven by a CMV promoter using the Adeno-Quest™ Kit (Quantum Biotechnologies, Inc., Montreal). A commercially available (Quantum Biotechnologies, Inc., Montreal) recombinant adenovirus  
 5 containing the beta-galactosidase gene was used as a control.

Initially, an *in vitro* dose response experiment was performed to determine the optimal concentration of adenovirus-delivered LMP-1 ("AdV-LMP-1") to induce bone differentiation in rat calvarial osteoblast cultures using a 60-minute transduction with a multiplicity of infection ("MOI") of 0.025, 0.25,  
 10 2.5, or 25 plaque-forming units (pfu) of virus per cell. Positive control cultures were differentiated by a 7-day exposure to  $10^{-9}$  M glucocorticoid ("GC"). Negative control cultures were left untreated. On day 14, the number of mineralized bone nodules was counted after von Kossa staining of the cultures, and the level of osteocalcin secreted into the medium (pmol/mL) was  
 15 measured by radioimmunoassay (mean  $\pm$  SEM).

The results of this experiment are shown in Table I. Essentially no spontaneous nodules formed in the untreated negative control cultures. The data show that a MOI equal to 0.25 pfu/cell is most effective for osteoinducing bone nodules, achieving a level comparable to the positive control (GC).  
 20 Lower and higher doses of adenovirus were less effective.

TABLE I

Outcome	AdV-LMP-1 Dose (MOI)					
	Neg. Ctrl.	GC	0.025	0.25	2.5	25
Bone Nodules	0.5 $\pm$ 0.2	188 $\pm$ 35	79.8 $\pm$ 13	145.1 $\pm$ 13	26.4 $\pm$ 15	87.6 $\pm$ 2
Osteocalcin	1.0 $\pm$ 0.1	57.8 $\pm$ 9	28.6 $\pm$ 11	22.8 $\pm$ 1	18.3 $\pm$ 3	26.0 $\pm$ 2

*In vivo* experiments were then performed to determine if the optimal *in vitro* dose was capable of promoting intertransverse process spine fusions in skeletally mature New Zealand white rabbits. Nine rabbits were anesthetized and 3 cc of bone marrow was aspirated from the distal femur through the intercondylar notch using an 18 gauge needle. The buffy coat was then isolated, a 10-minute transduction with AdV-LMP-1 was performed, and the cells were returned to the operating room for implantation. Single level posterolateral lumbar spine arthrodesis was performed with decortication of transverse processes and insertion of carrier (either rabbit devitalized bone matrix or a collagen sponge) containing 8-15 million autologous nucleated buffy coat cells transduced with either AdV-LMP-1 (MOI = 0.4) or AdV-BGal (MOI = 0.4). Rabbits were euthanized after 5 weeks and spine fusions were assessed by manual palpation, plain x-rays, CT scans, and undecalcified histology.

The spine fusion sites that received AdV-LMP-1 induced solid, continuous spine fusion masses in all nine rabbits. In contrast, the sites receiving AdV-BGal, or a lower dose of AdV-LMP-1 (MOI = 0.04) made little or no bone and resulted in spine fusion at a rate comparable to the carrier alone (< 40%). These results were consistent as evaluated by manual palpation, CT scan, and histology. Plain radiographs, however, sometimes overestimated the amount of bone that was present, especially in the control sites. LMP-1 cDNA delivery and bone induction was successful with both of the carrier materials tested. There was no evidence of systemic or local immune response to the adenovirus vector.

These data demonstrate consistent bone induction in a previously validated rabbit spine fusion model which is quite challenging. Furthermore, the protocol of using autogenous bone marrow cells with intraoperative *ex vivo* gene transduction (10 minutes) is a more clinically feasible procedure than other methods that call for overnight transduction or cell expansion for weeks in culture. In addition, the most effective dose of recombinant adenovirus (MOI=0.25) was substantially lower than doses reported in other gene therapy

applications (MOI-40-500). We believe this is due to the fact that LMP-1 is an intracellular signaling molecule and may have powerful signal amplification cascades. Moreover, the observation that the same concentration of AdV-LMP-1 that induced bone in cell culture was effective *in vivo* was also surprising given the usual required increase in dose of other growth factors when translating from cell culture to animal experiments. Taken together, these observations indicate that local gene therapy using adenovirus to deliver the LMP-1 cDNA is possible and the low dose required will likely minimize the negative effects of immune response to the adenovirus vector.

10

EXAMPLE 30: Use of Peripheral Venous Blood Nucleated Cells (Buffy Coat) for Gene Therapy With LMP-1 cDNA To Make Bone

In four rabbits we performed spine fusion surgery as above (Example 29) except the transduced cells were the buffy coat from venous blood rather than bone marrow. These cells were transfected with Adeno-LMP or pHIS-LMP plasmid and had equivalent successful results as when bone marrow cells were used. This discovery of using ordinary venous blood cells for gene delivery makes gene therapy more feasible clinically since it avoids painful marrow harvest under general anesthesia and yields two times more cells per mL of starting material.

20

EXAMPLE 31: Isolation of Human LMP-1 Splice Variants

Intron/Exon mRNA transcript splice variants are a relatively common regulatory mechanism in signal transduction and cellular/tissue development. Splice variants of various genes have been shown to alter protein-protein, protein-DNA, protein-RNA, and protein-substrate interactions. Splice variants may also control tissue specificity for gene expression allowing different forms (and therefore functions) to be expressed in various tissues. Splice variants are a common regulatory phenomenon in cells. It is possible that the LMP

25

30



splice variants may result in effects in other tissues such as nerve regeneration, muscle regeneration, or development of other tissues.

To screen a human heart cDNA library for splice variants of the HLMP-1 sequence, a pair of PCR primer corresponding to sections of SEQ ID NO: 22 was prepared. The forward PCR primer, which was synthesized using standard techniques, corresponds to nucleotides 35-54 of SEQ ID NO: 22. It has the following sequence:

5' GAGCCGGCATCATGGATTCC 3' (SEQ ID NO: 35)

10

The reverse PCR primer, which is the reverse complement of nucleotides 820-839 in SEQ ID NO: 22, has the following sequence:

5' GCTGCCTGCACAATGGAGGT 3' (SEQ ID NO: 36)

15

The forward and reverse PCR primers were used to screen human heart cDNA (ClonTech, Cat No. 7404-1) for sequences similar to HLMP-1 by standard techniques, using a cycling protocol of 94°C for 30 seconds, 64°C for 30 seconds, and 72°C for 1 minute, repeated 30 times and followed by a 10 minute incubation at 72°C. The amplification cDNA sequences were gel-purified and submitted to the Emory DNA Sequence Core Facility for sequencing. The clones were sequenced using standard techniques and the sequences were examined with PCGENE (Intelligenetics; Programs SEQUIN and NALIGN) to determine homology to SEQ ID NO: 22. Two homologous nucleotide sequences with putative alternative splice sites compared to SEQ ID NO: 22 were then translated to their respective protein products with Intelligenetic's program TRANSL.

20

25

One of these two novel human cDNA sequences (SEQ ID NO: 37) comprises 1456 bp:

30

10 20 30 40 50 60

CGACGCAGAG CAGCGCCCTG GCCGGGCCAA GCAGGAGCCG GCATCATGGA TTCCTTCAAG  
 70 80 90 100 110 120  
 GTAGTGCTGG AGGGGCCAGC ACCTTGGGGC TTCCGGCTGC AAGGGGGCAA GGACTTCAAT  
 130 140 150 160 170 180  
 5 GTGCCCCCTCT CCATTTCCCG GCTCACTCCT GGGGGCAAAG CGGCGCAGGC CGGAGTGGCC  
 190 200 210 220 230 240  
 GTGGGTGACT GGGTGCTGAG CATCGATGGC GAGAATGCGG GTAGCCTCAC ACACATCGAA  
 250 260 270 280 290 300  
 GCTCAGAACAA AGATCCGGGC CTGCGGGGAG CGCCTCAGCC TGGGCCTCAG CAGGGCCCAG  
 10 310 320 \* 330 340 \* 350 360  
 CCGGTTTCTG GCAAACCGCA GAAGGTGCAG ACCCCTGACA AACAGCCGCT CCGACCGCTG  
 370 380 390 400 410 420  
 GTCCCAGATG CCAGCAAGCA GCGGCTGATG GAGAACACAG AGGACTGGCG GCCGCGGCCG  
 430 440 450 460 470 480  
 15 GGGACAGGCC AGTCGCGTTC CTTCCGCATC CTTGCCCACC TCACAGGCAC CGAGTTCATG  
 490 500 510 520 530 540  
 CAAGACCCGG ATGAGGAGCA CCTGAAGAAA TCAAGCCAGG TGCCAGGAC AGAAGCCCCA  
 550 560 570 580 590 600  
 GCCCCAGCCT CATCTACACC CCAGGAGCCC TGGCCTGGCC CTACCGCCCC CAGCCCTACC  
 20 610 620 630 640 650 660  
 AGCCGCCCCG CCTGGGCTGT GGACCCTGCG TTTGCCGAGC GCTATGCCCC GGACAAAACG  
 670 680 690 700 710 720  
 AGCACAGTGC TGACCCGGCA CAGCCAGCCG GCCACGCCCA CGCCGCTGCA GAGCCGCACC  
 730 740 750 760 770 780  
 25 TCCATTGTGC AGGCAGCTGC CGGAGGGGTG CCAGGAGGGG GCAGCAACAA CGGCAAGACT  
 790 800 810 820 830 840  
 CCCGTGTGTC ACCAGTGCCA CAAGGTCATC CGGGGCCGCT ACCTGGTGCC GTTGGGCCAC  
 850 860 870 880 890 900

```

      GCGTACCACC CGGAGGAGTT TGTGTGTAGC CAGTGTGGGA AGGTCCTGGA AGAGGGTGGC
            910            920            930            940            950            960
      TTCTTTGAGG AGAAGGGCGC CATCTTCTGC CCACCATGCT ATGACGTGCG CTATGCACCC
            970            980            990            1000           1010           1020
5      AGCTGTGCCA AGTGCAAGAA GAAGATTACA GGCGAGATCA TGCACGCCCT GAAGATGACC
            1030           1040           1050           1060           1070           1080
      TGGCACGTGC ACTGCTTTAC CTGTGCTGCC TGCAAGACGC CCATCCGGAA CAGGGCCTTC
            1090           1100           1110           1120           1130           1140
      TACATGGAGG AGGGCGTGCC CTATTGCGAG CGAGACTATG AGAAGATGTT TGGCACGAAA
10      1150           1160           1170           1180           1190           1200
      TGCCATGGCT GTGACTTCAA GATCGACGCT GGGGACCGCT TCCTGGAGGC CCTGGGCTTC
            1210           1220           1230           1240           1250           1260
      AGCTGGCATG ACACCTGCTT CGTCTGTGCG ATATGTCAGA TCAACCTGGA AGGAAAGACC
            1270           1280           1290           1300           1310           1320
15      TTCTACTCCA AGAAGGACAG GCCTCTCTGC AAGAGCCATG CCTTCTCTCA TGTGTGAGCC
            1330           1340           1350           1360           1370           1380
      CCTTCTGCCC ACAGCTGCCG CGGTGGCCCC TAGCCTGAGG GGCCTGGAGT CGTGGCCCTG
            1390           1400           1410           1420           1430           1440
      CATTTCTGGG TAGGGCTGGC AATGGTTGCC TTAACCCTGG CTCCTGGCCC GAGCCTGGGC
20      1450
      TCCCGGGCCC TGCCCA

```

Reading frame shifts caused by the deletion of a 119 bp fragment  
(between \*) and the addition of a 17 bp fragment (underlined) results in a  
truncated gene product having the following derived amino acid sequence  
(SEQ ID NO: 38):

```

      10            20            30            40            50            60

```

```

MDSFKVVLEG PAPWGFRLQG GKDFNVPLSI SRLTPGGKAA QAGVAVGDWV LSIDGENAGS
      70          80          90          100          110          120
LTHIEAQNKI RACGERLSLG LSRAQPVQNK PQKVOTPDQK PLRPLVPDAS KQRLMENTED
      130          140          150          160          170          180
5  WRPRPGTGQS RSFRILAHLT GTEFMQDPDE EHLKKSSQVP RTEAPAPASS TPQEPWPGPT
      190          200          210          220          230          240
APSPTSPPW AVDPFAERY APDKTSTVLT RHSQPATPTP LQSRTSIVQA AAGGVPGGGS
      250          260          270          280          290          300
NNGKTPVCHQ CHQVIRARYL VALGHAYHPE EFVCSQCGKV LEEGGFFEEK GAIFCPPCYD
10      310          320          330          340          350          360
VRYAPSCAKC KKKITGEIMH ALKMTWHVLC FTCAACKTPI RNRAFYMEEG VPYCERDYEK
      370          380          390          400          410          420
MFGTKCQWCD FKIDAGDRFL EALGFSWHDY CFVCAICQIN LEGKTFYSKK DRPLCKSHAF
SHV

```

- 15 This 423 amino acid protein demonstrates 100% homology to the protein shown in Sequence ID No. 10, except for the sequence in the highlighted area (amino acids 94-99), which are due to the nucleotide changes depicted above.

The second novel human heart cDNA sequence (SEQ ID NO: 39)  
 20 comprises 1575 bp:

```

      10          20          30          40          50          60
CGACGCAGAG CAGCGCCCTG GCCGGGCCAA GCAGGAGCCG GCATCATGGA TTCCTTCAAG
      70          80          90          100          110          120
25 GTAGTGCTGG AGGGGCCAGC ACCTTGGGGC TTCCGGCTGC AAGGGGGCAA GGACTTCAAT
      130          140          150          160          170          180
GTGCCCCCTCT CCATTTCCCG GCTCACTCCT GGGGGCAAAG CGGCGCAGGC CGGAGTGGCC
      190          200          210          220          230          240

```

GTGGGTGACT GGGTGCTGAG CATCGATGGC GAGAATGCGG GTAGCCTCAC ACACATCGAA  
 250 260 270 280 290 300  
 GCTCAGAACAA AGATCCGGGC CTGCGGGGAG CGCCTCAGCC TGGGCCTCAG CAGGGCCCCAG  
 310 320 330 340 350 360  
 5 CCGGTTCAGA GCAAACCGCA GAAGGCCTCC GCGGACCTCC GCGGTACACC  
 370 380 390 400 410 420  
 TTTGCACCCA GCGTCTCCCT CAACAAGACG GCGGCGCCCT TTGGGGCGCC CCCGCCCCGT  
 430 440 450 460 470 480  
 GACAGCGCCC CGCAACAGAA TGGGTGCAGA CCCCTGACAA ACAGCCGCTC CGACCGCTGG  
 10 490 500 510 520 530 540  
 TCCCAGATGC CAGCAAGCAG CGGCTGATGG AGAACACAGA GGAAGTGGCGG CCGCGGCCGG  
 550 560 570 580 590 600  
 GGACAGGCCA GTCGCGTTCC TTCCGCATCC TTGCCCACCT CACAGGCACC GAGTTCATGC  
 610 620 630 640 650 660  
 15 AAGACCCGGA TGAGGAGCAC CTGAAGAAAT CAAGCCAGGT GCCCAGGACA GAAGCCCCAG  
 670 680 690 700 710 720  
 CCCCAGCCTC ATCTACACCC CAGGAGCCCT GGCCTGGCCC TACCGCCCCC AGCCCTACCA  
 730 740 750 760 770 780  
 GCGGCCCCGCC CTGGGCTGTG GACCCTGCGT TTGCCGAGCG CTATGCCCCG GACAAAACGA  
 20 790 800 810 820 830 840  
 GCACAGTGCT GACCCGGCAC AGCCAGCCGG CCACGCCAC GCGCTGCAG AGCCGCACCT  
 850 860 870 880 890 900  
 CCATTGTGCA GGCAGCTGCC GGAGGGGTGC CAGGAGGGGG CAGCAACAAC GGCAAGACTC  
 910 920 930 940 950 960  
 25 CCGTGTGTCA CCAGTGCCAC AAGGTCATCC GGGGCCGCTA CCTGGTGGCG TTGGGCCACG  
 970 980 990 1000 1010 1020  
 CGTACCACCC GGAGGAGTTT GTGTGTAGCC AGTGTGGGAA GGTCTGGAA GAGGGTGGCT  
 1030 1040 1050 1060 1070 1080

```

TCTTTGAGGA GAAGGGCGCC ATCTTCTGCC CACCATGCTA TGACGTGCGC TATGCACCCA
      1090      1100      1110      1120      1130      1140
GCTGTGCCAA GTGCAAGAAG AAGATTACAG GCGAGATCAT GCACGCCCTG AAGATGACCT
      1150      1160      1170      1180      1190      1200
5  GGCACGTGCA CTGCTTTACC TGTGCTGCCT GCAAGACGCC CATCCGGAAC AGGGCCTTCT
      1210      1220      1230      1240      1250      1260
ACATGGAGGA GGGCGTGCCC TATTGCGAGC GAGACTATGA GAAGATGTTT GGCACGAAAT
      1270      1280      1290      1300      1310      1320
GCCATGGCTG TGACTTCAAG ATCGACGCTG GGGACCGCTT CCTGGAGGCC CTGGGCTTCA
10      1330      1340      1350      1360      1370      1380
GCTGGCATGA CACCTGCTTC GTCTGTGCGA TATGTCAGAT CAACCTGGAA GGAAAGACCT
      1390      1400      1410      1420      1430      1440
TCTACTCCAA GAAGGACAGG CCTCTCTGCA AGAGCCATGC CTTCTCTCAT GTGTGAGCCC
      1450      1460      1470      1480      1490      1500
15 CTTCTGCCCA CAGCTGCCGC GGTGGCCCTT AGCCTGAGGG GCCTGGAGTC GTGGCCCTGC
      1510      1520      1530      1540      1550      1560
ATTTCTGGGT AGGGCTGGCA ATGGTTGCCT TAACCCTGGC TCCTGGCCCG AGCCTGGGCT
      1570
CCCGGGCCCT GCCCA

```

20        Reading frame shifts caused by the addition of a 17 bp fragment  
(bolded, italicized and underlined) results in an early translation stop codon at  
position 565-567 (underlined).

The derived amino acid sequence (SEQ ID NO: 40) consists of 153  
amino acids:

```

25
      10      20      30      40      50      60
MDSFKVVLEG PAPWGFRLQG GKDFNVPLSI SRLTPGGKAA QAGVAVGDWV LSIDGENAGS
      70      80      90      100     110     120

```

LTHIEAQNKI RACGERLSLG LSRAQPVQSK POKASAPAAD PPRYTFAPSV SLNKTARPPG

130

140

150

APPPADSAPO QNGCRPLTNS RSDRWSOMPA SSG

- 5 This protein demonstrates 100% homology to SEQ ID NO: 10 until amino acid 94, where the addition of the 17 bp fragment depicted in the nucleotide sequence results in a frame shift. Over amino acids 94-153, the protein is not homologous to SEQ ID NO: 10. Amino acids 154-457 in SEQ ID NO: 10 are not present due to the early stop codon depicted in nucleotide sequence.

EXAMPLE 32: Genomic HLMP-1 Nucleotide Sequence

Applicants have identified the genomic DNA sequence encoding HLMP-1, including putative regulatory elements associated with HLMP-1 expression.

- 5 The entire genomic sequence is shown in SEQ ID. NO: 41. This sequence was derived from AC023788 (clone RP11-564G9), Genome Sequencing Center, Washington University School of Medicine, St. Louis, MO.

- The putative promoter region for HLMP-1 spans nucleotides 2,660-8,733 in SEQ ID NO: 41. This region comprises, among other things, at least  
10 ten potential glucocorticoid response elements ("GREs") (nucleotides 6148-6153, 6226-6231, 6247-6252, 6336-6341, 6510-6515, 6552-6557, 6727-6732, 6752-6757, 7738-7743, and 8255-8260), twelve potential Sma-2 homologues to Mothers against Drosophilla decapentaplegic ("SMAD") binding element—sites (nucleotides 3569-3575, 4552-4558, 4582-4588, 5226-5232, 6228-6234,  
15 6649-6655, 6725-6731, 6930-6936, 7379-7384, 7738-7742, 8073-8079, and 8378-8384), and three TATA boxes (nucleotides 5910-5913, 6932-6935, and 7380-7383). The three TATA boxes, all of the GREs, and eight of the SMAD binding elements ("SBEs") are grouped in the region spanning nucleotides 5,841-8,733 in SEQ ID NO: 41. These regulatory elements can be used, for  
20 example, to regulate expression of exogenous nucleotide sequences encoding proteins involved in the process of bone formation. This would permit systemic administration of therapeutic factors or genes relating to bone formation and repair, as well as factors or genes associated with tissue differentiation and development.

- 25 In addition to the putative regulatory elements, 13 exons corresponding to the nucleotide sequence encoding HLMP-1 have been identified. These exons span the following nucleotides in SEQ ID NO: 41:

	Exon 1	8733-8767
30	Exon 2	9790-9895
	Exon 3	13635-13787



	Exon 4	13877-13907
	Exon 5	14387-14502
	Exon 6	15161-15297
	Exon 7	15401-15437
5	Exon 8	16483-16545
	Exon 9	16689-16923
	Exon 10	18068-18248
	Exon 11	22117-22240
	Exon 12	22323-22440
10	Exon 13	22575-22911

In HLMP-2 there is another exon (Exon 5A), which spans nucleotides 14887-14904.

All cited publications and patents are hereby incorporated by reference in their entirety.

15        While the foregoing specification teaches the principles of the present invention, with examples provided for the purpose of illustration, it will be appreciated by one skilled in the art from reading this disclosure that various changes in form and detail can be made without departing from the true scope of the invention.

We claim:

1. An isolated nucleic acid molecule comprising a SEQ ID NO: 37 or SEQ ID NO: 39.
2. An isolated human LMP protein encoded by SEQ ID NO: 37 or SEQ ID NO: 39.
3. A vector comprising the isolated nucleic acid molecule of claims 1.
4. A host cell comprising the vector of claim 3.
5. The host cell of claim 4, wherein the host cell is selected from the group consisting of prokaryotic cells, yeast cells and mammalian cells
6. The isolated nucleic acid molecule of claim 1, further comprising a label.
7. A human LIM mineralization protein comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 38 and SEQ ID NO: 40.
8. A monoclonal antibody specific for a HLMP-2 (SEQ ID NO: 38) or HLMP-3 (SEQ ID NO: 40).
9. A method of inducing bone formation comprising transfecting osteogenic precursor cells or peripheral blood leukocytes with an isolated nucleic acid molecule comprising SEQ ID NO: 39.
10. The method of claim 9, wherein the isolated nucleic acid molecule is in a vector.
11. The method of claim 10, wherein the vector is an expression vector.
12. The method of claim 11, wherein the vector is a plasmid.
13. The method of claim 11, wherein the vector is a virus.
14. The method of claim 13, wherein the virus is an adenovirus.
15. The method of claim 13, wherein the virus is a retrovirus.
16. The method of claim 9, wherein the osteogenic precursor cells or peripheral blood leukocytes are transfected *ex vivo*.
17. The method of claim 9, wherein the osteogenic precursor cells are transfected *in vivo* by direct injection of the isolated nucleic acid molecule.

18. The method of claim 9, wherein the LIM mineralization protein is HLMP-3 (SEQ ID NO: 40).

19. A method of fusing a spine, comprising:

- (a) transfecting osteogenic precursor cells or peripheral blood leukocytes with an isolated nucleic acid molecule comprising SEQ ID NO: 39;
  - (b) admixing the transfected osteogenic precursor cells or peripheral blood leukocytes with a matrix; and
  - (c) contacting the matrix with the spine;
- wherein expression of the nucleotide sequence causes mineralized bone formation in the matrix.

20. The method of claim 19, wherein the osteogenic precursor cells or peripheral blood cells are transfected *ex vivo*.

21. A method of inducing systemic bone formation in a mammalian host in need thereof, comprising:

- a) transfecting osteogenic precursor cells or peripheral blood leukocytes with a vector that is stably maintained in the osteogenic precursor cells or peripheral blood leukocytes, the vector comprising SEQ ID NO: 39 operably linked to a regulatable promoter, wherein the regulatable promoter responds to an exogenous control compound; and
- (b) administering to the host, as needed, an amount of the exogenous control substance effective to cause expression of SEQ ID NO: 39.

22. A method of stimulating production of an osteogenic soluble factor by an osteogenic cell, comprising:

(a) transfecting the osteogenic cell or peripheral blood leukocyte with an isolated nucleic acid molecule comprising SEQ ID NO: 39; and

5 (b) overexpressing the isolated nucleic acid molecule.

23. An osteogenic soluble factor produced by the method of claim 22.

24. The osteogenic soluble factor of claim 23, wherein the osteogenic factor is a protein.

25. A method of inhibiting the expression of HLMP-2 or HLMP-3  
10 comprising transfecting a cell wherein HLMP-2 or HLMP-3 is expressed with an antisense oligonucleotide.

26. The method of claim of claim 17, wherein the isolated nucleic acid molecule is in a vector selected from the group consisting of a plasmid and a virus.

15 27. The method of claim 26, wherein the vector is a plasmid, which plasmid is directly injected into muscle tissue.

28. A method of inhibiting bone formation comprising transfecting osteogenic precursor cells or peripheral blood leukocytes with an isolated nucleic acid molecule comprising SEQ ID NO: 37.

20 29. The method of claim 28, wherein the isolated nucleic acid molecule is in a vector.

30. The method of claim 29, wherein the vector is an expression vector.

31. The method of claim 30, wherein the vector is a plasmid.

32. The method of claim 30, wherein the vector is a virus.

25 33. The method of claim 32, wherein the virus is an adenovirus

34. The method of claim 32, wherein the virus is a retrovirus.

35. The method of claim 28, wherein the osteogenic precursor cells or peripheral blood leukocytes are transfected *ex vivo*.

30 36. The method of claim 28, wherein the osteogenic precursor cells are transfected *in vivo* by direct injection of the isolated nucleic acid molecule.

37. The method of claim 28, wherein the LIM mineralization protein is HLMP-2 (SEQ ID NO: 38).

## SEQUENCE LISTING

<110> Boden M.D., Scott D  
Hair Ph.D., Gregory A

<120> LIM Mineralization Protein Splice Variants

<130> splice variant LMP

<140> 6148.0222-00304

<141> 2000-04-28

<150> 60/132,021

<151> 1999-04-30

<160> 42

<170> PatentIn Ver. 2.1

<210> 1

<211> 457

<212> PRT

<213> Rattus norvegicus

<400> 1

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe  
1 5 10 15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg  
20 25 30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp  
35 40 45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile  
50 55 60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly  
65 70 75 80

Leu Ser Arg Ala Gln Pro Ala Gln Ser Lys Pro Gln Lys Ala Leu Thr  
85 90 95

Pro Pro Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Ala Ser Leu  
100 105 110

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Thr Asp Ser Ala  
115 120 125

Leu Ser Gln Asn Gly Gln Leu Leu Arg Gln Leu Val Pro Asp Ala Ser  
130 135 140

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly  
145 150 155 160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175  
 Glu Phe Met Gln Asp Pro Asp Glu Glu Phe Met Lys Lys Ser Ser Gln  
 180 185 190  
 Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Thr Ile Pro Gln Glu  
 195 200 205  
 Ser Trp Pro Gly Pro Thr Thr Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220  
 Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240  
 Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255  
 Asn Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Gly Thr Gly Gly  
 260 265 270  
 Gly Ser Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Lys Ile  
 275 280 285  
 Ile Arg Gly Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu  
 290 295 300  
 Glu Phe Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe  
 305 310 315 320  
 Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro Ser Cys Tyr Asp Val Arg  
 325 330 335  
 Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile  
 340 345 350  
 Met His Ala Leu Lys Met Thr Trp His Val Pro Cys Phe Thr Cys Ala  
 355 360 365  
 Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly  
 370 375 380  
 Ala Pro Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys  
 385 390 395 400  
 Arg Gly Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala  
 405 410 415  
 Leu Gly Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln  
 420 425 430  
 Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Lys Pro Leu  
 435 440 445  
 Cys Lys Ser His Ala Phe Ser His Val  
 450 455

<210> 2  
 <211> 1696  
 <212> DNA  
 <213> *Rattus norvegicus*

<400> 2  
 gcacgaggat cccagcgagg ctcttgagg cgcaggca gccgcccagc cgggcattca 60  
 ggagcaggta ccatggattc cttcaaggta gtgctggagg gacctgcccc ttggggcttc 120  
 cgtctgcaag ggggcaagga cttcaacgtg cccctctcca tctctcggct cactcctgga 180  
 ggcaaggccg cacaggccgg tgtggccgtg ggagactggg tactgagtat cgacggtgag 240  
 aacgcccga gacctacaca cattgaagcc cagaacaaga tccgtgcctg tggggagcgc 300  
 ctacgctgg gtcttagcag agcccagcct gctcagagca aaccacagaa ggccctgacc 360  
 cctcccgccg accccccgag gtacactttt gcaccaagcg cctccctcaa caagacggcc 420  
 cggcccttcg gggcaccccc acctactgac agcgccctgt cgcagaatgg acagctgctc 480  
 agacagctgg tccctgatgc cagcaagcag cggctgatgg agaatactga agactggcgc 540  
 ccgcgccag ggacaggcca gtcccgcttc ttccgcatcc ttgctcacct caggggcaca 600  
 gagttcatgc aagaccggga tgaggaattc atgaagaagt caagccagggt gccagggaca 660  
 gaagccccag cccagcctc aaccataccc caggaatcct ggccctggccc caccaccccc 720  
 agccccacca gccgcccacc ctgggcccgt gatcctgcat ttgctgagcg ctatgcccc 780  
 gacaaaacca gcacagtgt gaccgacac agccagccag ccacacctac gcctctgcag 840  
 aaccgcacct ccatagtcca ggctgcagct ggagggggca caggaggagg cagcaacaat 900  
 ggcaagacgc ctgtatgcca ccagtgccac aagatcatcc gccggcgata cctggtagca 960  
 ctggggccag cgtaccatcc tgaggaattt gtgtgcagcc agtgtgggaa ggtcctggaa 1020  
 gagggtggct tcttcgagga gaaggagct atcttttgcc cctcctgcta tgatgtgcgc 1080  
 tatgcaccca gctgtgcca atgcaagaag aagatcactg gagagatcat gcatgcgctg 1140  
 aagatgacct ggcattgttc ctgcttcacc tgtgcagcct gcaaaacccc tatccgcaac 1200  
 agggctttct acatggagga gggggctccc tactgcgagc gagattacga gaagatgttt 1260  
 ggcacaaagt gtcgcggtg tgacttcaag atcgatgccg gggaccgttt cctggaagcc 1320  
 ctgggtttca gctggcatga tacgtgtttt gtttgcgcaa tatgtcaaat caacttggaa 1380  
 ggaaagacct tctactccaa gaaggacaag cccctgtgca agagccatgc ctttccccc 1440  
 gtatgagcac ctctcacac tactgccacc ctactctgcc agaagggtga taaaatgaga 1500  
 gagctctctc tccctcgacc tttctgggtg gggctggcag ccattgtcct agccttggct 1560  
 cctggccaga tcctggggct cctcctcac agtcccctt cccacacttc ctccaccacc 1620  
 accacgtca ctacagggtg ctacgctcct agcccagtt cactctggtg tcacaataaa 1680  
 cctgtatgta gctgtg 1696

<210> 3  
 <211> 260  
 <212> DNA  
 <213> *Rattus norvegicus*

<400> 3  
 ttctacatgg aggagggggc tccctactgc gagcgagatt acgagaagat gtttggcaca 60  
 aagtgtcgcg gctgtgactt caagatcgat gccggggacc gtttcctgga agccctgggt 120  
 ttcagctggc atgatacgtg ttttgtttgc ccaatatgtc aaatcaactt ggaaggaaaag 180  
 accctctact ccaagaagga caagcccctg tgcaagagcc atgccttttc ccacgtatga 240  
 gcacctctc acactactgc 260

<210> 4  
 <211> 16  
 <212> DNA  
 <213> MMLV



<400> 4  
aagctttttt tttttg

16

<210> 5  
<211> 13  
<212> DNA  
<213> MMLV

<400> 5  
aagcttggt atg

13

<210> 6  
<211> 223  
<212> DNA  
<213> Homo sapiens

<400> 6  
atccttgctc acctcacggg caccgagttc atgcaagacc cggatgagga gcacctgaag 60  
aaatcaagcc aggtgcccag gacagaagcc ccagccccag cctcatctac accccaggag 120  
ccctggcctg gccctaccgc cccagccct accagccgcc cgcctgggc tgtggaccct 180  
gcgtttgccg agcgctatgc cccagacaaa accagcacag tgc 223

<210> 7  
<211> 717  
<212> DNA  
<213> Homo sapiens

<400> 7  
atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttcg gctgcaaggg 60  
ggcaaggact tcaatgtgcc cctctccatt tcccggctca ctctggggg caaagcggcg 120  
caggccggag tggcgtggg tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 180  
ctcacacaca tcgaagctca gaacaagatc cgggcctgcg gggagcgct cagcctgggc 240  
ctcagcaggg cccagccggt tcagagcaaa ccgcagaagg cctccgcccc cgcgcgggac 300  
cctccgcggt acacctttgc acccagcgtc tccctcaaca agacggcccc gccctttggg 360  
gcgccccgc cgcgtgacag cgcgccgcaa cagaatggac agccgctccg accgctggc 420  
ccagatgccg gcaagcagcg gctgatggag aacacagagg actggcgggc gcggccgggg 480  
acaggccagt cgcgttcctt ccgcatcctt gccacactca caggcaccga gttcatgcaa 540  
gaccgggatg aggagcacct gaagaaatca agccagggtc ccaggacaga agccccagcc 600  
ccagcctcat ctacacccca ggagccctgg cctggcccta ccgccccag cctaccagc 660  
cgccccctt gggctgtgga cctgcgttt gccgagcgt atgccccga caaacg 717

<210> 8  
<211> 1488  
<212> DNA  
<213> Homo sapiens

<400> 8  
atcgatggcg agaatgcggg tagcctcaca cacatcgaag ctgagaacaa gatccggggc 60  
tgccggggagc gcctcagcct gggcctcagc agggcccagc cggttcagag caaaccgcag 120  
aaggcctccg ccccgccgc ggaccctccg cggtagacct ttgcacccag cgtctccctc 180  
aacaagacgg cccggccctt tggggcgccc ccgcccgtg acagcgcccc gcaacagaat 240  
ggacagccgc tccgaccgct ggtcccagat gccagcaagc agcggtgat ggagaacaca 300

```

gaggactggc ggccgcggcc ggggacaggc cagtcgcgtt ccttccgcat ccttgcccac 360
ctcacaggca ccgagttcat gcaagaccgc gatgaggagc acctgaagaa atcaagccag 420
gtgcccagga cagaagcccc agccccagcc tcattctacac ccaggagcc ctggcctggc 480
cctaccgccc ccagccctac cagccgcccg ccctgagctg tggaccctgc gtttgccgag 540
cgctatgccc cggacaaaac gagcacagtg ctgaccgggc acagccagcc ggccacgccc 600
acgccgctgc agagccgcac ctccattgtg caggcagctg ccggaggggt gccaggaggg 660
ggcagcaaca acggcaagac tcccgtgtgt caccagtgcc acaaggatcat ccggggccgc 720
tacctggtgg cgttggggcca cgcgtaccac ccggaggagt ttgtgtgtag ccagtgtggg 780
aaggctcctg aagagggtgg cttctttgag gagaaggggc ccatcttctg cccaccatgc 840
tatgacgtgc gctatgcacc cagctgtgcc aagtgcagaa agaagattac aggcgagatc 900
atgcacgccc tgaagatgac ctggcacgtg cactgcttta cctgtgctgc ctgcaagacg 960
cccatccgga acagggcctt ctacatggag gagggcgtgc cctattgcga gcgagactat 1020
gagaagatgt ttggcacgaa atgccatggc tgtgacttca agatcgacgc tggggaccgc 1080
ttctggagg ccctgggctt cagctggcat gacacctgct tcgtctgtgc gatatgtcag 1140
atcaacctgg aaggaaagac cttctactcc aagaaggaca ggccctctctg caagagccat 1200
gccttctctc atgtgtgagc cccttctgcc cacagctgcc gcggtggccc ctacgctgag 1260
gggctggag tcgtggccct gcatttctgg gtagggtgg caatgggtgc cttaaccttg 1320
gtcctctggc cgagcctggg ctcccgggcc cctgccacc caccttatcc tcccaccca 1380
ctccctccac caccacagca caccggtgct ggccacacca gcccccttc acctccagt 1440
ccacaataaa cctgtaccca gctgaattcc aaaaaatcca aaaaaaaa 1488

```

&lt;210&gt; 9

&lt;211&gt; 1644

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

```

atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaaggg 60
ggcaaggact tcaatgtgcc cctctccatt tcccggtca ctctggggg caaagcggcg 120
caggccggag tggccgtggg tgactgggtg ctagcatcg atggcgagaa tgcgggtgag 180
ctcacacaca tcgaagctca gaacaagatc cgggcttgcg gggagcgct cagcctgggc 240
ctcagcaggg ccagccgggt tcagagcaaa ccgcagaagg cctccgcccc cgcccgggac 300
cctccgcggt acacctttgc acccagcgtc tccctcaaca agacggcccg gccctttggg 360
gcgccccgc cgctgacag cgccccgcaa cagaatggac agccgctccg accgctggtc 420
ccagatgcca gcaagcagcg gctgatggag aacacagagg actggcggcc gcggccgggg 480
acaggccagt cgcgttcctt ccgcattcct gccccacctc caggcaccga gttcatgcaa 540
gacccggtat aggagcacct gaagaaatca agccaggtgc ccaggacaga agccccagcc 600
ccagcctcat ctacacccca ggagccctgg cctggcccta ccgccccag ccctaccagc 660
cgcccgccct gggctgtgga ccctgcgttt gccgagcgct atgccccgga caaaacgagc 720
acagtgetga cccggcacag ccagccggcc acgcccacgc cgctgcagag ccgcacctcc 780
attgtgcagg cagctgccgg aggggtgcca ggagggggca gcaacaacgg caagactccc 840
gtgtgtcacc agtgccacaa ggtcatccgg ggccgctacc tggtggcgtt gggccacgcg 900
taccacccgg aggagtttgt gtgtagccag tgtgggaagg tcctggaaga ggggtggctt 960
tttgaggaga agggcgccat cttctgcca ccatgctatg acgtgcgcta tgcaccagc 1020
tgtgccaaat gcaagaagaa gattacaggc gagatcatgc acgcccagaa gatgacctgg 1080
cacgtgcaat gctttacctg tgctgcctgc aagacgcccc tccggaacag ggccttctac 1140
atggaggagg gcgtgccta ttgcgagcga gactatgaga agatgtttgg cacgaaatgc 1200
catggctgtg acttcaagat cgacgctggg gaccgcttcc tggaggccct gggcttcagc 1260
tggcatgaca cctgcttcgt ctgtgcgata tgctcagatc acctggaagg aaagaccttc 1320
tactccaaga aggacaggcc tctctgcaag agccatgcct tctctcatgt gtgagccct 1380
tctgcccaca gctgccggg tggcccttag cctgaggggc ctggagtcgt ggccctgcat 1440
ttctgggtag ggctggcaat ggttgcccta accctggctc ctggcccgag cctgggctcc 1500
cgggccctg cccaccacc ttatcctccc accccactcc ctccaccacc acagcacacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgcac aataaacctg taccagctg 1620
aattccaaaa aatccaaaa aaaa 1644

```

&lt;210&gt; 10

&lt;211&gt; 457

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe  
 1 5 10 15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg  
 20 25 30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp  
 35 40 45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile  
 50 55 60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly  
 65 70 75 80

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala  
 85 90 95

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu  
 100 105 110

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala  
 115 120 125

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser  
 130 135 140

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly  
 145 150 155 160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln  
 180 185 190

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu  
 195 200 205

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220

Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240

Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255

Ser Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly

260	265	270
Gly Ser Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Lys Val 275 280 285		
Ile Arg Gly Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu 290 295 300		
Glu Phe Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe 305 310 315 320		
Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro Pro Cys Tyr Asp Val Arg 325 330 335		
Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile 340 345 350		
Met His Ala Leu Lys Met Thr Trp His Val His Cys Phe Thr Cys Ala 355 360 365		
Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly 370 375 380		
Val Pro Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys 385 390 395 400		
His Gly Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala 405 410 415		
Leu Gly Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln 420 425 430		
Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Arg Pro Leu 435 440 445		
Cys Lys Ser His Ala Phe Ser His Val 450 455		

&lt;210&gt; 11

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 11

gccagggttt tcccagtcac ga

22

&lt;210&gt; 12

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 12

gccagggttt tcccagtcac ga

22

<210> 13  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 13  
tcttagcaga gcccagcctg ct 22

<210> 14  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 14  
gcatgaactc tgtgcccgtg ag 22

<210> 15  
<211> 20  
<212> DNA  
<213> Rattus norvegicus

<400> 15  
atccttgctc acctcacggg 20

<210> 16  
<211> 22  
<212> DNA  
<213> Rattus norvegicus

<400> 16  
gcactgtgct gggtttgtct gg 22

<210> 17  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 17  
catggattcc ttcaaggtag tgc 23

<210> 18  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 18  
gttttgtctg gggcagagcg 20

<210> 19  
<211> 44

<213> Artificial Sequence

<223> Description of Artificial Sequence: adaptor for  
Marathon RACE reactions

44

<213> Artificial Sequence

<223> Description of Artificial Sequence: PCR primer  
specific for Marathon RACE adaptor

27

<213> Homo sapiens

c c g t t g t t t g	t a a a a c g a c g	c a g a g c a g c g	c c c t g g c c g g	g c c a a g c a g g	a g c c g g c a t c	60
a t g g a t t c c t	t c a a g g t a g t	g c t g g a g g g g	c c a g c a c c t t	g g g g c t t c g	g c t g c a a g g g	120
g g c a a g g a c t	t c a a t g t g c c	c t c c t c c a t t	t c c c g g c t c a	c c t c t g g g g g	c a a g g c c g t g	180
c a g g c c g g a g	t g g c c g t a a g	t g a c t g g g t g	c t g a g c a t c g	a t g g c g a g a a	t g c g g g t a g c	240
c t c a c a c a c a	t c g a a g c t c a	g a a c a a g a t c	c g g g c c t g c g	g g g a g c g c c t	c a g c c t g g g c	300
c t c a a c a g g g	c c c a g c c g g t	t c a g a a c a a a	c c g c a a a a g g	c c t c c g c c c c	c g c c g c g g a c	360
c c t c c g c g g t	a c a c c t t t g c	a c c a a g c g t c	t c c c t c a a c a	a g a c g g c c c g	g c c c t t g g g g	420
g c g c c c c c g c	c c g c t g a c a g	c g c c c c g c a g	c a g a a t g g a c	a g c c g c t c c g	a c c g c t g g t c	480
c c a g a t g c c a	g c a a g c a g c g	g c t g a t g g a g	a a c a c a g a g g	a c t g g c g g g c	g c g g c c g g g g	540
a c a g g c c a g t	g c c g t t c c t t	t c g c a t c c t t	g c t c a c c t t a	c a g g c a c c g a	g t t c a t g c a a	600
g a c c c g g a t g	a g g a g c a c c t	g a a g a a a t c a	a g c c a g g t g c	c a g g a c a g a	a g c c c a g c c	660
c c a g c c t c a t	c t a c a c c c c a	g g a g c c c t g g	c c t g g c c c t a	c c g c c c c a g	c c c t a c c a g c	720
c g c c c g c c c t	g g g c t g t g g a	c c c t g c g t t t	g c c g a g c g c t	a t g c c		765

<213> Homo sapiens

cgacgcagag	cagcgccctg	gccggggccaa	gcaggagccg	gcattcatgga	ttccttcaag	60
gtagtgtctg	aggggccagc	accttggggc	ttccggctgc	aagggggcaa	ggacttcaat	120
gtgccctctc	ccattttccc	gctcaactct	ggggggcaaa	cgggcgcaggc	cggagtggac	180
gtgggtgact	gggtgtctag	catcgatgct	gagaatgcgg	gtagcctcac	acacatgga	240
gctcagaaca	agatccgggc	ctgcggggag	gcgctcagcc	tgggcctcag	cagggccccag	300

```

ccggttcaga gcaaaccgca gaaggcctcc gccccgcgcg cggaccctcc gcggtacacc 360
tttgaccca gcgtctccct caacaagacg gcccggccct ttggggcgcc cccgcccgt 420
gacagcgccc cgcaacagaa tggacagccg ctccgaccgc tgggccaga tggcagcaag 480
cagcggtga tggagaacac agaggactgg cggccgcggc cggggacagg ccagtccgt 540
tccttcgca tccttgcca cctcacaggc accgagttca tgcaagacc ggatgaggag 600
cacctgaaga aatcaagcca ggtgcccagg acagaagccc cagccccagc ctcatctaca 660
ccccaggagc cctggcctgg ccctaccgcc cccagcccta ccagccgcc gccctgggct 720
gtggaccctg cgtttgccga gcgctatgcc cgggacaaaa cgagcacagt gctgaccgg 780
cacagccagc cggccacgcc cagcccgctg cagagccgca cctccattgt gcaggcagct 840
gccggagggg tgccaggagg gggcagcaac aacggcaaga ctcccgtgtg tcaccagtgc 900
cacaaggtca tccggggccg ctacctggtg gcgttgggccc acgctacca cccgaggag 960
tttgtgtgta gccagtgtgg gaaggtcctg gaagagggtg gcttctttga ggagaagggc 1020
gccatcttct gccaccatg ctatgacgtg cgctatgcac ccagctgtgc caagtgaag 1080
aagaagatta caggcgagat catgcacgcc ctgaagatga cctggcacgt gcactgcttt 1140
acctgtgctg cctgcaagac gccatccgg aacagggcct tctacatgga ggagggcgtg 1200
ccctattgag agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagatcgacg ctgggggaccg cttcctggag gccctgggct tcagctggca tgacacctgc 1320
ttcgtctgtg cgatatgtca gatcaacctg gaaggaaaaga ccttctactc caagaaggac 1380
aggcctctct gcaagagcca tgccttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcctgga gtcgtggccc tgcatttctg ggtagggctg 1500
gcaatggttg ccttaacctt ggctcctggc ccgagcctgg gctcccgggc ccctgcccac 1560
ccaccttatc ctcccacccc actccctcca ccaccacagc acaccggtgc tggccacacc 1620
agcccccttt cacctccagt gccacaataa acctgtacct agctgaattc=caaaaaatcc 1680--
aaaaaaaaa 1689

```

```

<210> 23
<211> 22
<212> DNA
<213> Homo sapiens

```

```

<400> 23
gcactgtgct cgttttgtcc gg 22

```

```

<210> 24
<211> 21
<212> DNA
<213> Homo sapiens

```

```

<400> 24
tccttgcctca cctcacgggc a 21

```

```

<210> 25
<211> 30
<212> DNA
<213> Homo sapiens

```

```

<400> 25
tcctcatccg ggtcttgcac gaactcgggtg 30

```

```

<210> 26
<211> 28
<212> DNA

```

<213> Homo sapiens

<400> 26

gcccccgccc gctgacagcg ccccgcaa

28

<210> 27

<211> 24

<212> DNA

<213> Homo sapiens

<400> 27

tccttgctca cctcacgggc accg

24

<210> 28

<211> 22

<212> DNA

<213> Homo sapiens

<400> 28

gtaatacgac tcactatagg gc

22

<210> 29

<211> 23

<212> DNA

<213> Rattus norvegicus

<400> 29

gcggctgatg gagaatactg aag

23

<210> 30

<211> 23

<212> DNA

<213> Rattus norvegicus

<400> 30

atcttggtggc actggtggca tac

23

<210> 31

<211> 22

<212> DNA

<213> Rattus norvegicus

<400> 31

tgtgtcgggt cagcactgtg ct

22

<210> 32

<211> 1620

<212> DNA

<213> Homo sapiens



&lt;400&gt; 32

```

atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaaggg 60
ggcaaggact tcaatgtgcc cctctccatt tcccggctca ctctggggg caaagcggcg 120
caggccggag tggccgtggg tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 180
ctcacacaca tcgaagctca gaacaagatc cgggcctgcg gggagcgcct cagcctgggc 240
ctcagcaggg cccagccggg tcagagcaaa ccgcagaagg cctccgcccc cgcgcgggac 300
cctccgcggt acacctttgc acccagcgtc tccctcaaca agacggcccc gccctttggg 360
gcgccccgcg ccgctgacag cgccccgcaa cagaatggac agccgctccg accgctgggtc 420
ccagatgcca gcaagcagcg gctgatggag aacacagagg actggcggcc gcggccgggg 480
acaggccagt cgcgttcctt ccgcatcctt gccaccta caggcaccga gttcatgcaa 540
gaccgggatg aggagcacct gaagaaatca agccagggtg ccaggacaga agccccagcc 600
ccagcctcat ctacacccca ggagccctgg cctggcccta ccgccccag ccctaccagc 660
cgccccccct gagctgtgga ccctgcgttt gccgagcgct atgccccgga caaacagagc 720
acagtgtgta cccggcacag ccagccggcc acgcccacgc cgctgcagag ccgcacctcc 780
attgtgcagg cagctgccgg aggggtgcca ggagggggca gcaacaacgg caagactccc 840
gtgtgtcacc agtgccacaa ggtcatccgg ggccgctacc tgggtggcgtt gggccacgcg 900
taccaccggg aggagtgtgt gtgtagccag tgtgggaagg tcctggaaga ggggtggcttc 960
tttgaggaga agggcgccat cttctgcccc catgctatg acgtgcgcta tgcaccagc 1020
tgtgccaagt gcaagaagaa gattacaggc gagatcatgc acgcccctgaa gatgacctgg 1080
cacgtgcact gctttacctg tgctgcctgc aagacgcccc tccggaacag ggccttctac 1140
atggaggagg gcgtgcctta ttgcgagcga gactatgaga agatgtttgg cagcaaagtc 1200
catggctgtg acttcaagat cgacgctggg gaccgcttcc tggaggccct gggcttcagc 1260
tggcatgaca cctgcttcgt ctgtgcgata tgtcagatca acctggaagg aaagaccttc 1320
tactccaaga aggacaggcc tctctgcaag agccatgcct tctctcatgt gtgagccct 1380
tctgccaca gctgccgcgg tggcccctag cctgaggggc ctggagtcgt ggccctgcat 1440
ttctgggtag ggctggcaat ggttgccctta accctggctc ctggcccgag cctgggctcc 1500
cgggccccctg cccaccacc ttatcctccc acccaactcc ctccaccacc acagcacacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgccac aataaacctg taccagctg 1620

```

&lt;210&gt; 33

&lt;211&gt; 1665

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 33

```

cgacgcagag cagcgccctg gccggggcaa gcaggagccg gcatcatgga ttccttcaag 60
gtagtgtctg aggggcccag accttggggc tccggctgc aagggggcaa ggacttcaat 120
gtgccccctt ccatttcccg gctcactcct gggggcaaa cggcgaggc cggagtggcc 180
gtgggtgact ggggtgtgag catcgatggc gagaatgcg gtacccacac acacatcgaa 240
gctcagaaca agatccgggc ctgccccggg cgccctagcc tgggcctcag caggggccag 300
ccggttcaga gcaaaccgca gaaggcctcc gcccccgcc cggaacctcc gcggtacacc 360
tttgacacca gcgtctccct caacaagacg gcccgccct ttggggcgcc ccgccccgct 420
gacagcgccc cgcaacagaa tggacagccg ctccgaccgc tgggtcccaga tgccagcaag 480
cagcggctga tggagaacac agaggactgg cggcgcgggc cggggacagg ccagtgcgct 540
tccttccgca tccttgcccc cctcacaggc accgagttca tgcaagacc ggatgaggag 600
cacctgaaga aatcaagcca ggtgcccagg acagaagccc cagccccagc ctcatctaca 660
ccccaggagc cctggcctgg ccctaccgcc ccagcccta ccagccgccc gccctgagct 720
gtggaccctg cgtttgccga gcgctatgcc ccggacaaaa cgagcacagt gctgaccg 780
cacagccagc cggccacgcc cagcgcgtg cagagccgca cctccattgt gcaggcagct 840
gccggagggg tgccaggagg gggcagcaac aacggcaaga ctcccgtgtg tcaccagtgc 900
cacaaggtca tccggggccg ctacctggtg gcgttggggc acgcgtacca cccggaggag 960
tttggtgtga gccagtgtgg gaaggtcctg gaagagggtg gcttctttga ggagaagggc 1020
gccatcttct gccaccatg ctatgacgtg cgctatgcac ccagctgtgc caagtgcag 1080
aagaagatta caggcgagat catgcacgcc ctgaagatga cctggcacgt gcactgcttt 1140
acctgtgctg cctgcaagac gcccatccgg aacagggcct tctacatgga ggagggcggt 1200

```

```

ccctattgcg agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagatcgacg ctgggggaccg cttcctggag gccctgggct tcagctggca tgacacctgc 1320
ttcgtctgtg cgatatgtca gatcaacctg gaaggaaaga cttctactc caagaaggac 1380
aggcctctct gcaagagcca tgccttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcctgga gtcgtggccc tgcatttctg ggtagggctg 1500
gcaatggttg ccttaacctt ggctcctggc ccgagcctgg gctcccgggc ccctgcccac 1560
ccaccttata cttccacccc actccctcca ccaccacagc acaccggtgc tggccacacc 1620
agcccccttt cacctccagt gccacaataa acctgtaccc agctg 1665

```

&lt;210&gt; 34

&lt;211&gt; 223

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 34

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1             5             10             15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
          20             25             30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
      35             40             45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
      50             55             60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
      65             70             75             80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
          85             90             95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
          100             105             110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
      115             120             125

```

```

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser
      130             135             140

```

```

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly
      145             150             155             160

```

```

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr
          165             170             175

```

```

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln
          180             185             190

```

```

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu
          195             200             205

```

```

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro

```

210

215

220

<210> 35  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 35  
 gagccggcat catggattcc

20

<210> 36  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 gctgcctgca caatggagggt

20

<210> 37  
 <211> 1456  
 <212> DNA  
 <213> Homo sapiens

<400> 37  
 cgacgcagag cagcgccctg gccggggccaa gcaggagccg gcatcatgga ttccttcaag 60  
 gtagtgctgg agggggccagc accttggggc ttccggctgc aagggggcaa ggacttcaat 120  
 gtgcccctct ccatttcccg gctcactcct gggggcaaa cggcgcaggc cggagtggcc 180  
 gtgggtgact ggggtctgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240  
 gctcagaaca agatccgggc ctgcggggag cgcctcagcc tgggcctcag cagggcccag 300  
 ccggttcaga gcaaaccgca gaaggtgcag acccctgaca aacagccgct ccgaccgctg 360  
 gtcccagatg ccagcaagca ggggctgatg gagaacacag aggactggcg gccgcggccg 420  
 gggacaggcc agtcgcgttc cttccgcatc cttgcccacc tcacaggcac cgagttcatg 480  
 caagaccgag atgaggagca cctgaagaaa tcaagccagg tgcccaggac agaagcccca 540  
 gccccagcct catctacacc ccaggagccc tggcctggcc ctaccgcccc cagccctacc 600  
 agccgcccgc cctgggctgt ggaccctgcg tttgccgagc gctatgcccc ggacaaaacg 660  
 agcacagtgc tgacccgga cagccagccg gccacgcccc cgcgcgtgca gagccgcacc 720  
 tccattgtgc aggagctgc cggaggggtg ccaggagggg gcagcaacaa cggcaagact 780  
 cccgtgtgtc accagtgcga caaggtcatc cggggccgct acctggtggc gttggggcac 840  
 gcgtaccacc cggaggagtt tgtgtgtagc cagtgtggga aggtcctgga agaggtggc 900  
 ttctttgagg agaagggcgc catcttctgc ccaccatgct atgacgtgcg ctatgcaccc 960  
 agctgtgcca agtgcaagaa gaagattaca ggcgagatca tgcacgccct gaagatgacc 1020  
 tggcacgtgc actgctttac ctgtgctgcc tgcaagacgc ccatccgga cagggccttc 1080  
 tacatggagg agggcggtgc ctattgcgag cgagactatg agaagatgtt tggcacgaaa 1140  
 tgccatgggt gtgacttcaa gategacgct ggggaccgct tcctggaggc cctgggcttc 1200  
 agctggcatg acacctgctt cgtctgtgcg atatgtcaga tcaacctgga aggaaagacc 1260  
 ttctactcca agaaggacag gcctctctgc aagagccatg ccttctctca tgtgtgagcc 1320  
 ccttctgccc acagctgccg cgggtggccc tagcctgagg ggccctggagt cgtggccctg 1380  
 catttctggg tagggctggc aatggttgcc ttaaccctgg ctccctggccc gagcctgggc 1440  
 tcccgggccc tgccca 1456

<210> 38  
 <211> 423

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 38

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
 1             5             10             15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
          20             25             30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
          35             40             45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
          50             55             60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
 65             70             75             80

Leu Ser Arg Ala Gln Pro Val Gln Asn Lys Pro Gln Lys Val Gln Thr
          85             90             95

Pro Asp Lys Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser Lys Gln
100             105             110

Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly Thr Gly
115             120             125

Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr Glu Phe
130             135             140

Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln Val Pro
145             150             155             160

Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu Pro Trp
          165             170             175

Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp Ala Val
          180             185             190

Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser Thr Val
          195             200             205

Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln Ser Arg
210             215             220

Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly Gly Ser
225             230             235             240

Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Gln Val Ile Arg
          245             250             255

Ala Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu Glu Phe
          260             265             270

Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe Phe Glu

```

275                      280                      285  
 Glu Lys Gly Ala Ile Phe Cys Pro Pro Cys Tyr Asp Val Arg Tyr Ala  
 290                      295                      300  
 Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile Met His  
 305                      310                      315                      320  
 Ala Leu Lys Met Thr Trp His Val Leu Cys Phe Thr Cys Ala Ala Cys  
 325                      330                      335  
 Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly Val Pro  
 340                      345                      350  
 Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys Gln Trp  
 355                      360                      365  
 Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala Leu Gly  
 370                      375                      380  
 Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln Ile Asn  
 385                      390                      395                      400  
 Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Arg Pro Leu Cys Lys  
 405                      410                      415  
 Ser His Ala Phe Ser His Val  
 420

&lt;210&gt; 39

&lt;211&gt; 1575

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

cgacgcagag cagcgccctg gccgggcca gacagagccg gcatcatgga ttccttcaag 60  
 gtagtgctgg aggggcccag accttggggc ttccggctgc aagggggcaa ggacttcaat 120  
 gtgcccctct ccatctccc gctcactcct gggggcaaag cggcgagggc cggagtgagg 180  
 gtgggtgact ggggtgctgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240  
 gctcagaaca agatccgggc ctgcggggag cgcctcagcc tgggcctcag caggggcccag 300  
 ccggttcaga gcaaaccgca gaaggcctcc gcccccgccg cggaccctcc gcggtacacc 360  
 tttgcaccca gcgtctccct caacaagacg gcccggccct ttggggcgcc cccgcccgtc 420  
 gacagcgccc cgcaacagaa tgggtgcaga cccctgacaa acagccgctc cgaccgctgg 480  
 tcccagatgc cagcaagcag cggctgatgg agaacacaga ggactggcgg ccgcgggcgg 540  
 ggacagggca gtcgcttcc ttccgcatcc ttgcccacct cacaggcacc gagttcatgc 600  
 aagaccggga tgaggagcac ctgaagaaat caagccaggt gccaggaca gaagccccag 660  
 ccccagctc atctacaccc caggagccct ggccctggccc taccgcccc agccctacca 720  
 gccgcccgcc ctgggtgtg gacctgcgt ttgccgagcg ctatgcccc gacaaaacga 780  
 gcacagtgt gaccggcac agccagccgg ccacgcccac gccgctgcag agccgcacct 840  
 ccattgtgca ggcagctgcc ggaggggtgc caggaggggg cagcaacaac ggcaagactc 900  
 ccgtgtgtca ccagtgccac aaggtcatcc gggggccgta cctggtggcg ttggggccag 960  
 cgtaccaccc ggaggagttt gtgtgtagcc agtggtggaa ggtcctggaa gagggtggct 1020  
 tctttgagga gaaggcgcc atcttctgcc caccatgcta tgacgtgcgc tatgcacca 1080  
 gctgtgcaa gtgcaagaag aagattacag gcgagatcat gcacgcctg aagatgacct 1140  
 ggcacgtgca ctgctttacc tgtgctgcct gcaagacgcc catccggaac agggccttct 1200

```

acatggagga gggcgtgccc tattgcgagc gagactatga gaagatgttt ggcacgaaat 1260
gccatggctg tgacttcaag atcgacgctg gggaccgctt cctggaggcc ctgggcttca 1320
gctggcatga cacctgcttc gtctgtgcga tatgtcagat caacctggaa ggaaagacct 1380
tctactccaa gaaggacagg cctctctgca agagccatgc cttctctcat gtgtgagccc 1440
cttctgccc aagctgccc ggtggccct agcctgaggg gcctggagtc gtggccctgc 1500
atttctgggt agggctggca atggttgct taacctggc tcctggccc agcctgggct 1560
cccggccct gccca 1575

```

&lt;210&gt; 40

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1              5              10              15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
              20              25              30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
  35              40              45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
  50              55              60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
  65              70              75              80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
              85              90              95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
  100              105              110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
  115              120              125

```

```

Pro Gln Gln Asn Gly Cys Arg Pro Leu Thr Asn Ser Arg Ser Asp Arg
  130              135              140

```

```

Trp Ser Gln Met Pro Ala Ser Ser Gly
  145              150

```

&lt;210&gt; 41

&lt;211&gt; 24740

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 41

```

nnnnnntgta ttttatcata ttttaaaaat caaaaaacaa aaggcagttg aggttaggca 60
tgagggttcg tgacctgaat ccagcactt tgggaagccg aagcacgtgg atcacctgag 120
gtcaggagtt cgagaccagc ctgcccaata tggtaaaacc ctgtctctac taaaaataca 180
aaaaattagc caggcatggt ggtgggcacc tgtaatccca gctacttggg agactgaggc 240

```

```

aggagaatca cttaaacccg ggaggcgggc tgggcgcggt ggctcatgcc tgtaatccca 300
gcactttggg aggcgagac aggcgagatc tgaggtcagg agatcgagat catcctggct 360
aacatggtga aaccccatct ctactaaaaa tacaaaaaaa attagccagg cctggtggcg 420
ggcacctgta gtcccagcta cttgggagggc tgaggcagga gaatggcggtg aacctgggag 480
gcggcggttc agtgagccaa gatcgcgcca ctgcactcca gcctgggcca caagagtga 540
actccatctt aaagaaaaaa aacaaacccg ggaggcgga attgcagtca gccgagatct 600
cgccattgca ctcaagtatg ggtgacagag caagactcca tgtcaaaaaa aaaggcagtt 660
gacaggagca aggagcctgg tgaggaagct gtggcatttg acccggtgtg gttgctatgg 720
gccagggtgg tgctagtaga ggagctgagt gggaaagagc acaggggaca tgctgaaggc 780
ctgggtgtgg ggatgaggca gagattgggg gcaccttgca gggctcatagc aggtggctgt 840
ggtgagatgg aggaagacac ctgggggtact gctctaggct gtcagacata cagaagctgg 900
cccagccaag cccagggggt gcaagggaca tccttttgtg tccccagtga tctgcagctc 960
tcagacaccc tcaagcacag tgcctcttgc ccagcccagc actctcagtg gggagccagg 1020
tgggagaaca ggctcggaag gggacctagg cttatgcagc gagccgggca aagctggaac 1080
tgagagccag gccctggat gcccttggtc ttgtggagt ctgggatact gagggagg 1140
gacagggcag gggagtgcgg tgctctcacc tttagcttga actcattccc caggggacag 1200
gggaaacccc ctcaggatcc acagatgccc agtctcccaa gaggggctg gtcccatgg 1260
agggaaactc catctactcc tcctggcagg aaggttaagt ggaggacgtg caagggcagc 1320
ctcagcccc caccaccagg gctgggtctt ttgggactg acggagctgt cctggccacc 1380
tgccacagtg ggcgagtttc ccgtggtggt gcagaggact gaggcgcca cccgctgcca 1440
gctgaagggg ccggccctgc tgggtgctggg ccagacgccc atccagctga gggaggcca 1500
ggcaccagc ccctctacag ctggccctac cacttcctgc gcaagttcgg ctccgacaag 1560
gtgaggtgca ggggtgggaa agggtgaggg gctgacagcc tggaccctcc tgctaataccc 1620
caccctgtgt ccctgtgccc agggcgtgtt ctcctttgag gccggccgtc gctgccactc 1680
gggtgagggc ctctttgcct tcagcaccct ctgtgcccct gacctgtgca gggctgtggc 1740
cggggccatc gccgccagcg ggagcggctg ccagagctga ccaggcccca gccctgcccc 1800
ctgccacggg ccacctctct gccctccctg gacacccccg gagagcttcg ggagatgcca 1860
ccaggacctg agccacccac gtccaggaaa atgcacctgg ccgagcccg accccagagc 1920
ctgccgttac tgctaggccc ggagcccaac gatctggcgt ccgggctcta cgcttcagt 1980
tgcaagcgtg ccagtgggccc ccaggccaat gacacctct atgagaacct gtgtgtgctg 2040
gagccagccc ccacgctgca cgggtgggaa cctgagccgc acgagggccc cggagccgcg 2100
agccccacaa ccagtcccat ctaccacaac ggccaggact tgagctggcc cggcccgccc 2160
aacgacagta ccctggaggc ccagtaccgg cggctgctgg agctggatca ggtggagggc 2220
acaggccgcc ctgacctca ggcaggtttc aaggccaagc tggtgaccct gctgagtcgt 2280
gagcggagga agggcccagc cccttgtgac cggccctgaa cggccagcag agtggtggcc 2340
agaggggaga ggtgctcccc ctgggacagg agggtgggct ggtgggcaaa cattgggccc 2400
atgcagacac acgctgtgt ccacctggc ctgcaggaa aaggcaggcc gcctgtggag 2460
gacctcagcc ctgccctgcc ctctcatga atagtgtgca gactcacaga taataaagct 2520
cagagcagct cccggcaggg gactcacgg cacacgcccc tgcccacgtt cattgcccgc 2580
aacacaagca ccctgtgccg gttccagggg cacagggtgac ctgggcctta cctgccaccc 2640
gtgggctcaa acccactgca gcagacagac gggatggaaa tcattaggac tccatgttgc 2700
tctgcacggc cgagtgcac gaagaggcag gcggaggag ctgtgaggct tacttgtcag 2760
actcaggaag gagcaacatg agggcccaac tggagaccgc gagggccgag ctgggaggag 2820
gcagtggggg cgggtgagc gtggaaggga ttccagagac accctcgtcc aaaacacttg 2880
ttccctgctg aaactccaac aatttgaga tacttctggg aacccaggc gtcagtctcc 2940
tcatctgtaa aggagagaga accgatgacg tatcaggcat aatccttgat gagagtttgc 3000
tgctgcctta ctcagtgcc aagcgtgggg gacacagccg tggtcaggac agccttggtc 3060
ctgttctccg ggagccgaca ttccaggggg agagaagttt cctgaagact tccatgctgc 3120
gttccctct ctgctcctgc tcctggcgcc atcctaggag ccagccatgc acgcaagcgt 3180
catgcctcca gggctctgac tgcccagccc ctcaccgcaa ctcacctca gctgcacaca 3240
cccttggcac atcctgaacc tcattttcat gacggacaca caatttttgc tctctcctgt 3300
ccaagcctca tcctctggcc gccacctct tccagctcac ttcctttagt gcggccagta 3360
ccgcccctgc ctaggcatgt cgacctgcag ggacctttt ctggtcttc gaggcctctg 3420
cccaccatcc cctctttgtt ctccatagtc ccttccccct gttctctctc gtttcatctt 3480
actggtctgg caaagtcccc ggccttgggc gagccagacc tcctcagtgc ctgcacacag 3540
ctgcccacag ccagagaaat ccatttaagc agactgectg catccttctt aacagtgcaa 3600

```

```

ggcaggcact cctgccaca agagaccctg ttcctagta gggcagcttt tctcctcccc 3660
agaacctcct gtctatcccc acccaatgtc tctccacagg catattgggg aaacagggtca 3720
ggctctccca cgtatctgc aagtgtactg gcatccatct gtcttcttcc tacccttaca 3780
gtagaaacag tgtctgtccc cagctgtgct ctgatcccg ctcctttcac ctcagagctt 3840
ggaaaattga gctgtcccca ctctctcctg cgccattca tctaccage agcttttcca 3900
gccacacgca aacatgtctt gtaatttcac attttaaacc ttccttgac ctcacattcc 3960
tcttcggcca cctctgtttc tctgttcttc ttcacagcaa aaactgttca aaagagtgtg 4020
tgattacttt catttccact ttctcaccct cattctctcc tcaattaact ctccttcac 4080
cccatgatgc cattatgtgg cttttattag agtcaccaac cttattctcc aaaacaaaag 4140
caacaaggac tttgacttct cagcagcact cagctctggg tcttgaaaca ccccgttac 4200
ttgctattcc tcctacctca taacaatctc cttcccagcc tctactgctg ctttctctga 4260
gttcttccca gggctcctagg ctcagatgta gtgtagctca accctgctac acaaagaatc 4320
tcttgaaagc ctgtaaaaaat gtccatgcat gttctgtgag tgatctacca agaaaataaa 4380
aaattttaaa aatcaaatgc ccatgcctgg gccacacgc aggggctctg atttcatcag 4440
tctggtaggt ggggttctggg catccacgct cactggattt ccgatgatt gtagtatgca 4500
gcctaggctg ggaaccactg gcctcagcaa gccagtcatt ctcagggtgt cacagacct 4560
ctagggtgcta atgaccccgga aggtctgtct cagtgcaca cctccccctg agtccagat 4620
ttaggaatcc cactgcacac gagacatctg gatgtggaaa agacatctcc agatcccatg 4680
ggtgaaaggg ggttggggga atggagactc gtgttcttcc aggatgtgtg tggacacaga 4740
atgcaaagcc tggagggatg ctagagccat agggaggaag atttcggctc acttattcat 4800
gcaagcactt cctgatgggt aaggtcttag agcaagctga ggccaagagg cgggcagtcg 4860
aggtgctgct gcaggcacc cactcccta cagtggcaag cccaagcca gcccttggca 4920
gctcaaatcc caggacacgc tgaaggtcac ccagagagtc aggggcatgg ctagaaccag 4980
aaccaggac tctggggacc cagcatggca tcctttcctt cattacaaat ctgagctgct 5040
ttgtttccta gggatttctg tgatattcca aggggactgt gggaaagaaa gtccttggaa 5100
accaccagga cgctagaggc ctggcctgga gcctcaggag tctcgccac cagaggcgc 5160
tggttctctg tccagggtcca gttgctacgc aggggctgcc tgtgctggga ggctccccag 5220
gggacacaga ccagagcctt gcaccagccc aaggaatggg agcctggggg cctctctgct 5280
ggaggactgc caggaccccc aggtgcgc ctcttcttct gctcatttgc tgtttcactt 5340
tgtcaatcct tccttcttct gtgtgttcat tcacatccac tgtgtgctgg ccctggggaa 5400
atgttagata agacacatta gctgtgtgtc ttcattgtcc taacaaagaa cacaccctgg 5460
aaagagcacc gcagagagtc cccattcccc catctccctc cacacatgga atctggagat 5520
gccttttcca catccagatg tctctggtgc tgtgggattc ttaaataaac aaacatttca 5580
tacagaatgt gagatgatgg agatgctatg gggaaaagta aagcagaggg agggcctagt 5640
gtgtgatgcy ggtgaggcat ccagggattg ctgtttcagc tgtgatcagg aaaggcctg 5700
ggaggaggcc acatctgagc agagacctaa ataaagtgg aaacctgtg ctgagatct 5760
tggaagagtg tttcaagggc cgggcaccgg gcatggtggc tcacgcctgt aatcccagca 5820
ctttgggagg ccaaggcagg tggatcgctg gaggtcagga gtttgagagc agcctgacca 5880
acatggagaa accccatctc tactaaacat ataaaaatta tccgggcatg gtggttcatg 5940
cctgtagtcc cagctactcg ggaggttag gcaggagaat cacttgaacg tgggaggcag 6000
aggttgacgc aagccgagat cacaccactg cactccagcc tggatgacag agcgagactc 6060
cgtctcaaaa aaaaaaaga aaagaaaaaa gaaaaaaa gaaaagtgtt tcaagcaggg 6120
gaactggcaa gtggagaggc cctgaggcag aaatatgctt ggctgctgg aggaaatgtg 6180
agtgaggagg tcagggtggc tggagtggag ggagcgagt gtaggagtca gaccagttt 6240
attcatattc tgtaggtctt aaggacttca gttttatttt gagtgaata tgagccact 6300
ggaatgctaa aagctgagag tgacatggtg ctgtgattct ggcttataaa atatcacttt 6360
ggctgcttcg tgaagactct ggaaggggca agggtgaaag cagggatgcc cgttaggaga 6420
ccgttacagg ggcgcaggca caaatggca gtggctggga caatggtggc agcagcgggt 6480
agatgtgaac atgttgaagg tggaaatttgc agaactcggg ggaggacaga agagaaagga 6540
taacttcacg gtttctgctg aaccagtgg ataaatgttg gtggcacttc ttgaagttag 6600
gaaggagtta ggaaggtggg aaaggcacia gtttgaattg ggccatgatg gtctgagata 6660
cctagtacag tggttcccca accttttgg cagaagggac cgctttcatg gaagacaatt 6720
ttccacaga ctgggggtgg ggtggggatg gtttcagggt gggtcgagt cagtacattt 6780
atcattagac tctttttttt tttttttttt tgagatagag tctcgctctg tccccacac 6840
tggagtgcag tggagccatc ttggctcact acaacctctg ctgccagggt tcaagtcat 6900
ctcctgcctc agcctctcaa gtagctggga ttataggcat atgcgccacc acgccagct 6960

```



```

aattttttagta ttttttagtag agacgggggt tcaccatatt ggccaggatg gtctcgaact 7020
cctgacctca agtgatcctc ccccgccca acctcccaaa gtgctgggggt tacaggcgtg 7080
aaccactgca cccggcccat ttatcattag attctcataa ggaatgagca acctagatcc 7140
ctcgcatgca cagttcacaa tagggttcac gtcctatgg gagtctaag ctgccgtgc 7200
actcagcttc tctggcttgc cgctgctcac cttctgctgt gcagcccagt tccaaacagg 7260
ccacaaacgg ggagttgggg acccctgac tagtaaaca ctaggcaggg ttttgataa 7320
tggagttaga gttcctgggg agaggtcagg ctggccatga aacatgggat gcctttgcat 7380
ataggtgggtg ttgaaagcca caggacagta cggggtctca gggggtgagc ataaagagag 7440
gcgacatcag atggccaagg ccagaggcag aggaggatgg gaaggagggg ccagtggggc 7500
aggggggaagc tgtgaagcca gggaaaaagg gtgtttcgcg gaaaaggatc aacctggacc 7560
agtgtgccc ctaggcaggg caggatgaaa cttaaccacc acggattcca tggcccatg 7620
gcctccaggc cacaggggac cttgagaaga gagatctcag gggacgggtg cggacaagag 7680
cccgctggc atggctcaa gagataactg aaggaaagca agtgagacg cgataaacag 7740
acaactccct ggaggaattt tactctcag aggagaatta aaggtagta gctggagagg 7800
gatgtgggt caagagaagg tctttaacga cgagaactc cagcgcggtt tgtgcagaac 7860
agggtgggtg tgatgactgt ggatggagag gggagaactg cagcgactct gtcctaggag 7920
gaggtgatgg gccgggacca ccaagcaggt ggagggtgga cggcccttcc ctcacccga 7980
caccgcagtg tgcctagtgt ccgtgcgcgc gccctagtgt cctgggctga acgcgggg 8040
gggactctga ggacgcctcc caggcgcgca gtccgtctgg ccaagtgga gcgggacggc 8100
ngcttccgac ggtgcgcggg tgggctcggg gttgcaggga catccggcgt ccgctcctgc 8160
cctgttttcc tgccttcgca gagcgttgc caactctagc tttaaacgcc cctgtcccc 8220
tcaacttgct tccccagcc cctctgattt acagattctg cagtccccga gggttgcgc 8280
tacgataccg aactcgcg cagcctgcga ggcgagtatg atcgtcccat ttttcggagt 8340
agcaaaactaa ggttcagaga ctactatgtc ccaggctcgt ctggtttgaa ggtccgctt 8400
cctctccctc cgccagcggg cgggtgcgag gactgggcga ggacgcgtt ccctaaggag 8460
gcgacccgca gccccggccc cctcccgact ccgccccgtt gcagggcccg ggtcggcgag 8520
gcctctcagc tctaagcccg acgggacttg gtgattgggc aggacggaag agctgggtgg 8580
ggctttccac cagcggagaa agtctagtgg gcgtggtcgc gacgagggcg tggcctgggt 8640
ccccgcccc gtccgcgcgc tcaaagtga ggggtggtgt gggggcgggg tcagaacact 8700
ggcgccgat ccaacgagg ctccctggag cccgacgcag agcagcgccc tggccggg 8760
aagcaggtat cgacgaccgc gggggcgctc ttgggtgga ccaggcgggc gcccgggg 8820
tgctgaggac cacaaggggc actgggggtc gtggtccagg ctgtgcttcc tccgctggc 8880
cctggccctt gcctccgccc ccgccccgc cttcctgcg ctaagccggc tgcggcggg 8940
ccgattggcg cctgcgggt tctgcgcgc gggccagtct aatgcatggg gcccgggcg 9000
gggactaagg ggaactgag tcacgtcgg gtgggagcag ttctgtgtgg gaggcaccac 9060
ccccactgg gtcggggaa ggatccccct ccaagctatg ctgaggggtc ccagcccca 9120
tctgtctcca caggggcgcc acccactcc cgccttcccc ttcttcagca ccagggggtc 9180
ccgccctggc tcccagcagc ctcgactggt cccggaatgg ctagggagat ccgctgcagc 9240
cgctccctc ccctccctc ccctccctc ccctccctc ccctccctc ccctccctc 9300
cccctcggt ccaagcccc cgtgtgctc ctcgctggc tctccgcaca gtgtcagctt 9360
acacgcctta tatagtccga gcaggctcca gccgcggcct gctgccggga cctggggcg 9420
ggggagagga gagccggccc ctgactcacc cggaccgccc gaggtccag gctggcttgg 9480
ggggaggccg cgccagtta gtccctcgcc caccctgg ttgcaaagaa cctcaagcct 9540
ggattcaggc accctcacc gttccagtcc caaggggagg ggggctgctc ctgtcttcc 9600
aaagtgaagt ccgacagcca gcagcccagg ccagcctgac aaaatacctg cctcctatgg 9660
cttgggctg ctcaggggt gccgctgctt gcctggcccc tgtccaaggc tggatcctg 9720
agctggcccc gcctgcctgc ctgcccgc accatgctgg ccaactacct tctcttctct 9780
cctctcagga gccgcatca tggattcctt caaagtatg ctggaggggc cagcaccttg 9840
gggcttccgg ctgcaagggg gcaaggactt caatgtgccc ctctccattt cccgggtgag 9900
cctaggtttg gggagggggc tccccagcg gtctttcggt gcttaggtct ccagaggggtg 9960
atggggggag tccaaacagg agctggtcag gggccagcag gccaggagat gtctagggtcc 10020
ggagatgtag tggtaacctg ctgccacaag gactcccaat gaggtggata ctgggaggga 10080
gcacccaggc ttctccagcc ctgactgta cccgatgctg ttctcccaag ctctgtggc 10140
cacctctgag ggtggaggg aggtcattg tgcaggatgg gagcctaaca tttcaggagg 10200
tatctaaact tgaggtggca atgcttgag ccaggcccca ggcaggacac tgtgactata 10260
ggatttcact tcagcctcac tgccgcccag ggaatagcaa tctcatccc gttttccag 10320

```

```

atgagagaag aactcatgga gaggtggcgg ggctcgctca tcgagtccat ggtgaagcag 10380
ggattggaat tgaggcacag catggcgtag attttttggt ggtagaaggg gtctctcccc 10440
agcctatgta aggaccaca tccactgttc ccattcagga tgtggtggcc tttgacccca 10500
agcagaagtg taggacaggg ctccattcta ggggcttaac ttcagcttcc aagagcctgc 10560
cctggtgtgg gtggagctgg aggctggctc ctccctgtag cagggggatt gccttataag 10620
cccaagaatg cagccccacg ctgggatggc caacagtggc tgcggtctgc agagctgaaa 10680
agggctggcc taggcctggc cccctgaacc ccactggtgg gcctctcagc tggtcaccag 10740
gctgcagctc cagctgtatg gtccagtgtg gagacacaac aaattgcctg cccagagtgg 10800
gtgaggccag cctgtcggct ggcatctctg actggcctgg gggtcaggag ggggtgggga 10860
cttcctgccc ctatatccgc ctgccccgag agaccacccc aggcgcgggg tgggcaggca 10920
gctgttgtca ggaagcccaa ggcaagccca gcctggaggg gccagaggg tcgtggcctg 10980
aggaggggct caagctggag tctgtctgta ggagctgggc gtgggggtta ggggtggcag 11040
gccagcagtg ctcttctcag gggctccttg atggcattct cctggaacct gccccgccag 11100
cagggtagtg aggagtggt tgcctatga cacagctccc actacatagc cctcacacag 11160
ccctgaaacc tacctgacgt cctgctccct gggaaagtgc tggcccagtg tgtctgggga 11220
gcctgaacct cagtttcttc cctgatggag atgactttca gatatggcct gttgggggca 11280
ctccgggctc cagctccctg gtcagcatcc ctggcatgtg ggccggggcca ctagctgac 11340
ccagccctgg agttggacct gggccacat ggggtgggtga ggtgggcttt tctgagttag 11400
gccagcccc tccccctccc ctgaccccag aatggagggg ggtgggaggg gcaagggctg 11460
gctgtgggccc caggcctggg agatgaggta acgtctggga ctggggggct gggctgctca 11520
ggctgactca cccccacct atgcagggtc cagccccctg gctttttccc tcttggttc 11580
ctctggcctt accctgcccc tggcttgagc cctccctgc ctctctccag ccaccgccc 11640
agcgtgtct tctgctctcc tgcctgccc cccacgctct gaacaccct catcctctgt 11700
gcttcctgcc ctctcactc tgggaaggga agcgtcccc gccccccacc cctctccag 11760
gagccagcta gctgcacccc aagaccccca cctcgggctc agcccacagc tcccaggagc 11820
cagccctgtg ggcagggagt ggctgggcca ggtttccctt ctactgactc accatgacct 11880
tgagtaagtc acttcccctc tggggtgtca cttcccata cacagtataa ggggttgatt 11940
tagttggatt gaactaaagg tgagggagtg gctcagggtg tctccaggtg ggtgacccc 12000
tcagttgggc ccccatgctc agcagaggtg gccacagtg gtggagcctt agggtcagag 12060
acacttcctg gctctgctc ttactagctg ggtgacttga ggcaagttgt ttaacctctc 12120
tgtgtacatt tgcaagtgca aaatgggtaa aatcccagat tactccacaa ggttggtgga 12180
agattcagtg tcaatatgta gcatagttgg tgctcaataa actgaagcaa gtcttcttat 12240
ttagcgagtg aggaaggggc cgccagctc tcttagcctt ctgacctcct acgcaagcaa 12300
gaggtcatgt tgagcccagc tgcctttct tttcccagtg ctgtcaagct ctgtgcctgg 12360
ctgccctgcc ctctgacatc tctctgaaac ctcttgctc ccctctccct gcctcagctc 12420
agtctgtgca ctgaccacc tgaggagcct cctggggcca ctggcagcct ggaccccccc 12480
agatcccccc caccagtgga aattgtcttc cagcactgcc tcacaaaagc ctacttgatg 12540
cagtgccagg cctcttgcca gatggctggg tggctccctta ggcttgagc cagtcaagct 12600
gccctgctg tgttgctggg gctgggctag aggcctggaa ggggtttatc agggtcaccc 12660
tctcagggcc tgggagatac ccaatcccag acattaaaac tgccagtagc cctctacct 12720
tcaaagccaa gtcttggtcc ctcccttggt cattcaaagc catcgtaagt gaactctcac 12780
ccgctaggca gcacagcca ttctccttta ccgagggcca ccgcttctc aaagtcattc 12840
ctgatggctc cagctcatgc tgggtggcagc catttctccc agcctactgt ctctactcat 12900
tgccacagga accagggact cccagctcaa gagcctgaag gattggggtc aggggaaatt 12960
ggcagtcgag ggcttgggag tgacagccat gtatggccta cgaagtccca gctgtcaact 13020
taggtcccat tcaggcagtg ttcacaggga accgggagat aacagggcct gttcctggct 13080
ctcaaagggc cccagcagac ccctatagat ggccccgac aggggtgctg ggggtgagag 13140
gtccataaga gccccggtg gtttcgggga ggaagctgcc ccctgcatgg gccagagggc 13200
atatctggta ggtggagtgg cctgggcagg aggcagcag gagcctcaa aggcaatggt 13260
cctctgaaa cacttgggt ttagcctgag cgtggctgtt tgtggacatc atagcaattt 13320
ctggactgtg ggggaggggt gtggcggtga atagataagc atcgtgactg gggaagctca 13380
ggtgagcacc acctgaggga gagggctctg cagtgaataa ataagcagtg tgactgggaa 13440
attgtgaagc tcaggtgagc gccaccacct cctgggttgc tttagtgtcc agcagctgcc 13500
tagaactatg ttgaatgaag agctctctgg gttctggaag tgggacagct ttgggtgggg 13560
cagtgttacc accgtcagcc tggcttgggt ctgcagggtc cagggcctcg gtcactttgc 13620
ttctctctcc acagctcact cctgggggca aagcggcgca ggcggagtg gccgtgggtg 13680

```

actgggtgct	gagcatcgat	ggcgagaatg	cgggtagcct	cacacacatc	gaagctcaga	13740
acaagatccg	ggcctgcggg	gagcgctca	gcctgggcct	cagcaggat	gcgggtggac	13800
atggatgggt	gcgcccgcgc	tggcagtggg	gatccctgcg	ccccggcccg	ctgtcacgct	13860
ttccttctcc	tccagggtccc	agccggttca	gagcaaaccg	cagaaggtag	gaggctggcc	13920
gggacatccg	ggcgggtggg	ggtgtgggct	tggacggcca	ggcctgctcg	ccctcctggc	13980
acattctcgg	taccccaatc	cctggccggg	agtggagggc	agaaaccgga	gctaaggcgg	14040
gtctaggggc	ctggagtga	gccaggggct	gctgcacggg	cctggcacca	cgcatgtccg	14100
cctgtctgtc	cgctgtctg	tccgcctgct	gcctcccgc	gccggcgctg	cgtgctcgcc	14160
cgcactcggt	cagccctcgg	tctgtcgtgg	actgagatcg	ccactcccaa	atggggccct	14220
tgaaacctga	gtcgtcctct	ccccgtagcc	tccaaataga	tgtagggggg	gggggtgggg	14280
tggggggctg	gagctgccgc	tgtcctctgc	tgcaggcgcc	ccacttccac	ccaggccccc	14340
accttaccct	gcccgcggc	cctgcccggc	tgtgtctctg	cccaggcctc	cgcccccgcc	14400
gcggaccctc	cgcggtacac	ctttgcaccc	agcgtctccc	tcaacaagac	ggcccggcct	14460
ttgggcgccc	ccgcccgtg	acagcgcccc	gcagcagaat	gggtacgtcg	gcccctgccc	14520
gcccgcgccc	acgccatcag	gcccactgtg	gccccacgcc	cgctgcccgc	tgtgtctcag	14580
tctgtgctgc	gccccagccc	ggcggaaccg	tgcggcacgc	cccctggcgg	ccggggtggg	14640
gctgcaggca	cagggtccct	cccgaggctg	tggcgccctg	cagggtaccg	cctggggagg	14700
ggtctctgaa	tgacgcgcg	ccccctgctg	gcggctgggg	gttgggttgt	ggtgtcgggc	14760
cagctgagcc	ccagacactc	agtgcgcct	tgtcccgggc	tgttctgacc	cctcccgtc	14820
tttcttctc	tctgtgtct	gtccctttgt	ccctttatct	gtctgtctgt	cttatttct	14880
tcacagggtg	agaccctga	caagtcagt	agccccctc	tgctgtgccc	tttcttctc	14940
cttttggcac	tctgggtggc	ggccccctcc	cacctggct	gccctcctct	ccacttcgcc	15000
ctcctgtcct	ctcacctacc	cgcccagcag	ggctcctggc	ctcacctta	cccactccct	15060
cccactactg	taaccctaac	ccacatgcac	caaatcctgg	gaggggctgc	ccccaccgcc	15120
cacccccagt	gtgggttct	gagccacacc	ctccccacag	acagccgctc	cgaccgctgg	15180
tcccagatgc	cagcaagcag	cggtgatgg	agaacacaga	ggactggcgg	ccgcccggcg	15240
ggacaggcca	gtcgcgttcc	tccgcaccc	ttgcccacct	cacaggcacc	gagttcagta	15300
agtgccagcc	cagggcaggg	ggtactttcc	tgcccccag	cccaggcgtg	atccctgacc	15360
ctgtgtcttt	tttggtaaat	gcctgctct	gctctctcag	tgcaagacct	ggatgaggag	15420
cacctgaaga	aatcaaggta	cagggacggg	caccagcccc	tctccacct	cctgcctctt	15480
ccattccagc	tactgccctg	tgtctactcc	tgaggctccc	agctggggct	ctcaattctc	15540
ccttccctcc	ttccttccct	ccttccctcc	ttccttccct	ccttccctcc	ttccttccct	15600
cccttccctc	ttccttccct	cttccatttc	ttccctccct	ccttccctcc	ctcctccctc	15660
cctgcctccc	ttccatctct	ccttccctcc	acttcttcc	ccctctctct	ctgcccctca	15720
gggaaaagta	tgtcctggag	ctgcagagcc	cacgtacac	ccgctccgg	gactggcacc	15780
accagcgctc	tgcccacgtg	ctcaacgtgc	agtcgtagcc	cggccctctc	cagccggctg	15840
ccctctctgc	ctccctcttt	ctgttccctc	tgcccagggc	acccccctag	tgccctccagc	15900
ttctgcctac	ctcaccctcc	ctttcgtgcc	cctggcctga	gcctcctgct	ggcctggccc	15960
tggccgcccc	cctgggttca	tctgacactg	ccttccctct	ttgcctgtg	gtactgctgt	16020
ctgccagggtc	tgtgctgcct	tgggcatgga	ataaacattc	tcagccctgc	ttgctctgcc	16080
tgtcttctat	ctttgtggac	ctggtttgca	tttggggtgt	gggggtgttt	cgtggttcgg	16140
actgtttggg	ccctgcctgc	cttgttttca	gtgggagggg	gtacctggca	aaggggcct	16200
gccctgccat	cacagatggc	ttcctggcat	gaggggagcc	ccaggagctg	cctcagaagc	16260
gggagccctg	cctcgtctcc	cagctagaga	ccgcacacca	gctaactgga	cattgctagg	16320
agaagctgcc	cttcccatcc	ctacccaggt	gggacctgga	atccaactcg	gcagtttcca	16380
cgccccaggt	catctccctg	ggggccagca	ggaccaggt	tggggggtgg	ggccatgtca	16440
ggaagctcag	ccatgcaggg	ccttgaatgg	cagatcttgc	agccagggtg	ccaggacaga	16500
agccccagcc	ccagcctcat	ctacacccca	ggagccctgg	cctggtgaga	gggagtgggc	16560
tcgggcctgg	gcaagggtgg	gcagcctcca	ggggcatggg	ggtggtgggc	ttctctcagc	16620
tgctgggggc	tcaccccccg	tcctttgggg	tccttgggca	cccctttaga	gtcactttcc	16680
ccggcaggcc	ctaccgcccc	cagccctacc	agccgcccgc	cctgggctgt	ggaccctgcy	16740
tttgccgagc	gctatgcccc	ggacaaaacg	agcacagtgc	tgaccgggca	cagccagccg	16800
gccacgcccc	cgccgctgca	gagccgcacc	tccattgtgc	aggcagctgc	cggagggtgt	16860
ccaggagggg	gcagcaacaa	cggcaagact	cccgtgtgtc	accagtgcc	caaggtcacc	16920
cgggtgggtg	cctgttctct	tccgacctg	gctttcccat	cctgcagccc	agccccacct	16980
gtctgcccac	ctgtcttgcc	tcagctgcga	ctggggggaa	taaggattca	gttctcagct	17040

```

ggagtaggag tagggacctg ggctgggtcc tcccattctt aatcccacgc tacctacccc 17100
agcccaccca caacaactgc tagcagcatc tgccgtggcg aaatagccga agggccaacc 17160
ataggctgaa gctgcacccc tacctttgct gctctctggg caaagagggg cctgccccct 17220
cccagcgctg ctgccccctc ctctgtctct ctgtctccct ctgctctcag agcatacagg 17280
cctggagcca ctccctctgt gcactgcccc gtggggccaa gcagcatcaa acacccccca 17340
gcatacagct gccggattct agagccttcc taattgcag gcctggcctg ctctcatctc 17400
tgtcagctct tttttttttt tttttgaaac agagtctcac tgtgttggcc acgttggcgt 17460
gcagtggcgc gatctcggct cactgcaacc tctgcctcct gggttcaaga gattctcctg 17520
cctcagcctc ctgagtagct gggattacag gcacccgcca ccatgcctgg ctaattttgt 17580
atTTTTtagta gagacggggt tttaccatgt tggccaggct ggtctcaaac tctcacctc 17640
aggtgatctc aggcctgcct tggcctccca aagtgtctgg actacagggt tgagccactg 17700
tgcccagccg actctatcag ctcttgccag gtagaacagg caggccagca ggacagggca 17760
gctccagggt ttgcccaggg ggggtcagc ttttatgagg ctccagtcgt cageccttcc 17820
tcccggggct ctccctgctc taaagctgcc tctcctgtca ccagcagttc agtgtggcgg 17880
actggctctg taagcttcat ggctgccacg gtcacttccc aagcctgtct tctatcctat 17940
gtggaaaaat gggagaatga actgtccctc ccaaggcctc ctggtgggtg gtcagtcac 18000
ctgaaggggg ccaagacccc cactctctg cgtgtgctcc ctctgaccgc tctgcctccc 18060
ctgcaggggg cgctacctgg tggcgctggg ccacgcgtac caccggagg agtttgtgtg 18120
tagccagtgt gggaggtcc tggagagggt tggcttcttt gaggagaagg gcgccatctt 18180
ctgcccacca tgctatgacg tgcgtatgc acccagctgt gccaagtgc agaagaagat 18240
tacaggcgtg agtagggctg gctggcgggg aggtgggtccc aagcctgtca gtgggaacga 18300
gggtgctggt gaaaccacac gtccaggtct ctccccaggt gagcctccgg gtccttacca 18360
gcgtaataaa tgggtgctg tactggcctc accctgcatt agtcaggatg ctcttaacaa 18420
atgaccatgt tctgtctcag aaaccgcca aggtgcaaaa gagcaggagg accaagccag 18480
gagaagccct gggccctcct gactccact ttgggtctct cctgccctgg tgaaatgaca 18540
gaacggccaa cttgacacgc tgaagctgct ctgtctcatg cgtcctcctc atttctggat 18600
ccagagccag ggctgccagg agtagccaga gagctctgtg tgggtgatgt catattagt 18660
aggtttacct tgaccacgag cagtgggaaa ctcaaaataa tgggtggtta tttctcatct 18720
aaaaacatcc cgggtgggtt ggtctgggac tgatctggtg gaccagggt ccgccttgtt 18780
gcttgactgt tggcagcacc tgcttactta ccactcatgg tgcaagatga cacttcagcc 18840
tccgccaaaa tgtcacctt ccagccagca ggaagtcgga aggagaagaa aggggacaga 18900
gccccatggc gtccatcctt agaggatgct gccacctgaa cctctgcttt catcctgtt 18960
gtcagaaccc agtcacatga ccacaccag tggcaacgga ggctgggaaa tatagtcttt 19020
atTTTgggca cccatgtgtc cagcaaaact ggggggtcca tcagtcggca agaacgggag 19080
agtggcagat gcagtggctg atgcttgtat ccagcactt tgggagggtc aggtgggcag 19140
atcacctgag gtcaggagt caagaccagc ctggccaata tggtgaaacc ctgtctctac 19200
taaaaaataa aaaattagct ggggtgtgct ggcacactgt agtcccagct acttgggagg 19260
ctgaggcagg agaatcgct gatcttgaga ggtggagggt gcagttagcc aagattgtgc 19320
cactgccttc cagcctggga gacagcaaaa aaaaaaaaaa aaaaaaaaaa aaaaagggcc 19380
aggcacgggt gctcacacct gtaatcccag cactttggga ggccgagatg ggcgatcac 19440
gaggtcagga gattgagacc atcctggcta acacggtgaa accccatctc tactaaaaat 19500
acaaaaaaat tggccgggca tgggtggagta gtcccagcta ctggggaggc tgaggcagga 19560
gaatggcgtg aacctgggag gcagagcttg cagttagccg agatcgcgcc actgcactcc 19620
agcctgggca acagagcgag actcttgtct caaaaagaaa aaaagaaaga gaaatctgcc 19680
tcccagcctt gggctcctgc cctaccagcc cacaccctg gtagagcctc ctctcccacc 19740
agctcaaagc ccaagttcct tcaactgtgac cttgtctgct cctctaaaac aggcaacacc 19800
agacagttag aagagccagc cagacatggg cagaaaacct atttctgtga tctactggct 19860
gtgtgagcag gggctagtgt ctctctctgg gcctcactga agagaagggt ggcactatgc 19920
tagggccggc acggttgcaa ggtagatgta agatggggta cagggtgtgt ggagggcaga 19980
aatgcacat ccgaaggcta catgtcccc acacttatgt cttgcttggc ccacactgtt 20040
tcattttaaa atcagtagca aacaatttaa aaaatcagaa gatttgctg catgatgcag 20100
tggctcatgc ctgtaatccc agcactttgg gaggccaagg tgggaggatt gcttgagccc 20160
aggagttaa gaccagcatg ggcacatag caagaccct gtttctacaa aaaaaaaaaa 20220
attagaaaat tagccaagt tgggtggcatg cacctgtggt cccagctact tgggaggcag 20280
agggaaagt agatctcct ctttttatt ctttatgtat aatgatagg tcttgcctg 20340
ttgcccaggc tggagtgcag tggcatgat actgctcact gcagccttga tctcctgggc 20400

```

tcagaggatc	ctccccacctc	agcctcccaa	atagctagga	ctagagggtgc	ccaccagcat	20460
gctcagcaga	tttttaaatc	ttttttaga	gatgaggttt	tgctatggtg	cccaggctgg	20520
tctcgaactc	ctggcctcga	gcgatcctcc	caccttggcc	tcccaaagca	ctgggattac	20580
agacgtgagc	cactgcgccc	agcagatttc	tctttaacac	ctagatttca	gcctgagcca	20640
ggcaggcatt	cctgaatgaa	ccagtagtac	tgctcccaga	agaagaggtc	ctcctccgtg	20700
tgacacagtc	cccacttggc	ccttgcaggg	attggatctg	ggatccctgg	atttaaactc	20760
agggccatcc	tcataacagc	ctcacaaggc	tgggattagc	ttcccagttc	acaagggag	20820
aaaccaagac	ttgagaaggt	caaggtctgg	ccagaccac	acatcttgga	ccctcatacc	20880
gcctcgaggc	cccatgctgc	cctctgcctg	ctccagatgt	gaatactgct	ggccctggct	20940
ggccccggct	ggccccgagg	gtcctagggg	tgaacagccc	agcccaggga	gagctcagcc	21000
ccttgtgcct	ctgccccttc	ccacctcctg	cggaggccag	tcgactcacc	cacaaagggc	21060
caggcactgt	ggggatagat	cagctaacaa	aacagttgat	gcttctctgc	cttctggggc	21120
ttacattttg	gctggaagaa	gaggggagag	gcagactgta	agcaataagc	gcaataagta	21180
ggttgccctg	aagtaatggt	agatcacgtt	acggaaaaca	ggaaagagca	gagcgacaag	21240
tgctgggggtg	cgtggtgcag	ggaaggcagc	tggctgctgc	tggtgtgggtc	agagtgggccc	21300
ctcatggaga	agactgcatt	cgagcagaaa	cttgaagggg	gtgaggggtg	agcctagaga	21360
tatctggggc	agagcagtc	aggcagagg	gacagccgtg	gtcaagccca	ggacaggagt	21420
gtgcctgggtg	tgccagtttc	aggcaagagg	ccagtgtgca	gaggcaagg	gagaacgcaa	21480
gggagagcag	tgccggagac	gggtgggaac	gaggtcagac	ctgctggcct	ccagcctctg	21540
catggggcct	ggctcttgct	gggagcaatg	ggaagcagta	cacagtttca	tgcaagggga	21600
gaaggccctgt	cttgggttgc	aggggcaagc	tgtggcagct	gggatcagag	agaggagctt	21660
gtaggccagt	tggtatgtgg	tcccacgggc	cagatggcca	tggcttacct	cacttcaggg	21720
aggctgtgag	aagcactcag	aatctggatg	tgccttgggg	gtgggcccga	ctggatttcc	21780
tggtggacct	ggtgtggggg	gtgagaggag	ggtgtgtttg	gctgcagcag	acaggagaat	21840
ggagttgcc	tccgcgtgat	ggggatggct	gtgggaggag	aggtttgggg	tgagggaatc	21900
aggaactgag	tgctggacat	ggcaagtctg	aaggcgcagt	ggtcgtccac	tcagagacct	21960
tggagttgga	gatggagggtg	tgggagtcct	gaacagttag	atgtagtggt	taccgcgaga	22020
aggaacaggg	cttgcggcca	gccctcctgt	gttcccgtga	cccagggcag	ggcaggaggg	22080
gcctgagcct	gccgagtgac	tgggacctcc	ttccaggaga	tcatgcacgc	cctgaagatg	22140
acctggcacg	tgactgctt	tacctgtgct	gcctgcaaga	cgcccatccg	gaacagggcc	22200
ttctacatgg	aggagggcgt	gccctattgc	gagcgaggta	cccactggcc	agtgaggggtg	22260
aggagggatg	gtgcatgggg	caggcatgaa	tccaggtcct	ctttctctct	gccccattc	22320
tcagactatg	agaagatggt	tggcacgaaa	tgccatggct	gtgacttcaa	gatcgacgct	22380
ggggaccgct	tcctggaggc	cctgggcttc	agctggcatg	acacctgctt	cgtctgtgcg	22440
gtgagagccc	cgccccctga	actgagcccc	aagcccaccg	gccctctggt	cattccccag	22500
gagatgcagg	agaagttggg	aaggggcctc	tcctgctgcc	cccaacccca	tgtagactggg	22560
cctttgctgt	ccttagatat	gtcagatcaa	cctggaagga	aagaccttct	actccaagaa	22620
ggacaggcct	ctctgcaaga	gccatgcctt	ctctcatgtg	tgagccccct	ctgcccacag	22680
ctgccgcggt	ggccccctagc	ctgagggggc	tggagtctgt	gccctgcatt	tctgggtagg	22740
gctggcaatg	gttgccttaa	ccctggctcc	tggcccggagc	ctggggctcc	ctgggcccctg	22800
ccccaccac	cttatcctcc	caccccaactc	cctccaccac	cacagcacac	cgatgctggc	22860
cacaccagcc	ccctttcacc	tccagtgcc	caataaacct	gtaccagct	gtgtcttctg	22920
tgcccttccc	ctgtgcaccc	ggaggggcag	aatttgaggc	acgtggcagg	gtggagagta	22980
agatggtttt	cttgggctgg	ccatctgggt	ggtcctcgtg	atgcagacat	ggcgggctca	23040
tggttagtgg	aggaggtaca	ggcgagacc	catgtgccag	gcccgggtgcc	cacagacatg	23100
aggggagcca	ctggtctggc	ctggccttgg	ggttagagaa	gggtagttag	gaagggtagt	23160
tagcatgggtg	gctcatgcct	gtgatccag	cactttggaa	ggccaagggtg	ggcagatcgc	23220
ttgaggtcag	gagttcgaga	cctcatggcc	aacacggtga	aacagcgtct	ctagtaaaaa	23280
tacaaaaatt	agccgagtgt	ggtggggcat	gcctgtaatc	ccagccactc	aggaggctga	23340
ggcgggaaaa	tcacttgaac	ctgggaagtg	gaggttgag	tgagctgaga	tcacaccact	23400
gcgcgcgagc	ctgggtggca	gatggcagag	cgagaccctg	cttcaaaaaa	aaaaaaaaaa	23460
aaaaaaaaaa	gaagggtagt	tgtagtggg	ggtggatctg	cagagatatg	gtgtggaaaa	23520
cagcaatggc	cacagcaaag	tcttgagggg	gccagctgcc	gtccaaacag	aagaaggcag	23580
ggctggagag	ggtagccctt	aggtcctggg	aagccacgag	tgccaggcag	tagagctggg	23640
gctgtctctt	gaggttaggg	cagggcaagg	cacagcagag	tttgaaatag	gtttgtgttg	23700
tattgcagaa	aagaggcccc	agaacactga	gggagtgcag	gagggaggct	gggaggagga	23760

```
gttgccagcag ggcctagggg cggggggccag gcaagggagg ggcagagagt aatatggcag 23820
agatgggacc cagtggcagg tccgggggat gagggatgga gagaaggaca ggagcgttgc 23880
caggcatctg gcctatacca gacatgctca cgctgtctcc cgcgaaacct ctagcaacct 23940
tgcgccgttg tctgcaatca cttatttcat tttttctttt ttaactttaa ttttttttgt 24000
ttttaagaga caggatctcc ctaggttgcc cgggctggtt tcaaactcct gggctcaagc 24060
aattcttctt ccttagcccc aaagtgtctg cattacagg gtgagccacc atgcttgcc 24120
cacttatttt ctagatgagg cacagaaaga ttgggagact tgaccaaggt cacgctgtca 24180
ttgagccatg agccagacta gaatccaggc ctgaagctgg gtgcgctgtc ccaggactgg 24240
ctggcactga gtaccatttg ccagcgagca tctctctggg aagctgactt ctgcccggta 24300
cctggaggac tgtagacctt ggtggtggcg ccgtcactct ggggcttcct gcctcccact 24360
gatgcccgca ccaccctaga gggactgtca tctctcctgt cccaagcctg gactggaaag 24420
actgaagaga agccttaagt aggccaggac agctcagtgt gccatggctg cccgtccttc 24480
agtgggtccct ggcattgagga cctgcaacac atctgttagt cttctcaaca ggcccttggc 24540
ccggtccccct ttaagagacg agaagggtg ggcacggtga ctcacacctc taatcccagc 24600
actttggaag gctgaggctg gagaagggt ccagcttagg agttcaggac cagcctgggc 24660
aacatggtga gacctgttt tgttttgttt tttgtttttt tgagatggag tcttgctctg 24720
tcgcccaggc tggagtgcag 24740
```

<210> 42

<211> 25

<212> DNA

<213> Rattus norvegicus

<400> 42

gcactacctt gaaggaatcc atggt

25

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/11664

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : 435/320.1, 325; 530/350, 388.1; 536/23.5; 514/44

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 325; 530/350, 388.1; 536/23.5; 514/44

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BODEN et al. LMP-1, a LIM-domain protein, mediates BMP-6 effects on bone formation. Endocrinology. 1998. Vol. 139, No. 12, pages 5125-5134, entire document.	1-37
A	HOGAN et al. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes and Development. 1996. Vol. 10, pages 1580-1594, entire document.	1-37
A	KINGSLEY, DM What do BMPs do in mammals? Clues from the mouse short-ear mutation. Trends Genet. January 1994. Vol. 10, No. 1, pages 16-21, entire document.	1-37

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

16 JULY 2000

Date of mailing of the international search report

29 AUG 2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

ANNE-MARIE BAKER

Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US00/11664

**A. CLASSIFICATION OF SUBJECT MATTER:**

IPC (7):

A61K 48/00; C07H 21/04; C07K 14/00, 16/00; C12N 15/63, 15/85, 15/86

**B. FIELDS SEARCHED**

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST

Dialog (file: medicine)

search terms: gene(w)transfer, LMP, bone, LIM, mineral?, human, osteogen?



CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
9 November 2000 (09.11.2000)

PCT

(10) International Publication Number  
WO 00/66178 A1

(51) International Patent Classification<sup>7</sup>: A61K 48/00,  
C07H 21/04, C07K 14/00, 16/00, C12N 15/63, 15/85,  
15/86

(21) International Application Number: PCT/US00/11664

(22) International Filing Date: 28 April 2000 (28.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/132,021 30 April 1999 (30.04.1999) US

(71) Applicant (for all designated States except US): EMORY  
UNIVERSITY [US/US]; 1380 South Oxford Road, At-  
lanta, GA 30322 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BODEN, Scott, D.  
[US/US]; 2842 Cravey Drive, N.E., Atlanta, GA 30345  
(US). HAIR, Gregory, A. [US/US]; 1887 Chrysler Drive,  
N.E., Atlanta, GA 30345 (US).

(74) Agents: GARRETT, Arthur, S. et al.; Finnegan, Hen-  
derson, Farabow, Garrett & Dunner, L.L.P., 1300 I Street,  
N.W., Washington, DC 20005-3315 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE,  
DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,  
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,  
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ,  
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT,  
TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent  
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent  
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,  
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,  
GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- With international search report.
- With (an) indication(s) in relation to deposited biological  
material furnished under Rule 13bis separately from the  
description.

(48) Date of publication of this corrected version:

15 March 2001

(15) Information about Correction:

see PCT Gazette No. 11/2001 of 15 March 2001, Section  
II

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

WO 00/66178 A1

(54) Title: LIM MINERALIZATION PROTEIN SPLICE VARIANTS

(57) Abstract: The present invention is directed to isolated nucleic acid molecules that encode LIM mineralization protein, or LMP. The invention further provides vectors comprising splice variants of nucleotide sequences that encode LMP, as well as host cells comprising those vectors. Moreover, the present invention relates to methods of inducing bone formation by transfecting osteogenic precursor cells with an isolated nucleic acid molecule comprising a nucleotide sequence encoding splice variants of LIM mineralization protein. The transfection may occur *ex vivo* or *in vivo* by direct injection of virus or naked plasmid DNA. In a particular embodiment, the invention provides a method of fusing a spine by transfecting osteogenic precursor cells with an isolated nucleic acid molecule having a nucleotide sequence encoding LIM mineralization protein, admixing the transfected osteogenic precursor cells with a matrix and contacting the matrix with the spine. Finally, the invention relates to methods for inducing systemic bone formation by stable transfection of host cells with the vectors of the invention.

## LIM Mineralization Protein Splice Variants

## BACKGROUND OF THE INVENTION

## 5      1. Field of the Invention

The field of the invention relates generally to osteogenic cells and the formation of bone and boney tissue in mammalian species. Specifically, the invention concerns a novel family of proteins, and nucleic acids encoding those proteins, that enhances the efficacy of bone mineralization *in vitro* and *in vivo*.

10      The invention provides methods for treating a variety of pathological conditions associated with bone and boney tissue, such as, for example, spine fusion, fracture repair and osteoporosis.

## 2. Description of the Related Art

15      Osteoblasts are thought to differentiate from pluripotent mesenchymal stem cells. The maturation of an osteoblast results in the secretion of an extracellular matrix which can mineralize and form bone. The regulation of this complex process is not well understood but is thought to involve a group of signaling glycoproteins known as bone morphogenetic proteins (BMPs). These  
20      proteins have been shown to be involved with embryonic dorsal-ventral patterning, limb bud development, and fracture repair in adult animals. B. L. Hogan, Genes & Develop., 10:1580 (1996). This group of transforming growth factor-beta superfamily secreted proteins has a spectrum of activities in a variety of cell types at different stages of differentiation; differences in  
25      physiological activity between these closely related molecules have not been clarified. D. M. Kingsley, Trends Genet., 10:16 (1994).

To better discern the unique physiological role of different BMP signaling proteins, we recently compared the potency of BMP-6 with that of BMP-2 and BMP-4, for inducing rat calvarial osteoblast differentiation. Boden *et al.*,  
30      Endocrinology, 137:3401 (1996). We studied this process in first passage (secondary) cultures of fetal rat calvaria that require BMP or glucocorticoid for

initiation of differentiation. In this model of membranous bone formation, glucocorticoid (GC) or a BMP will initiate differentiation to mineralized bone nodules capable of secreting osteocalcin, the osteoblast-specific protein. This secondary culture system is distinct from primary rat osteoblast cultures which  
5 undergo spontaneous differentiation. In this secondary system, glucocorticoid resulted in a ten-fold induction of BMP-6 mRNA and protein expression which was responsible for the enhancement of osteoblast differentiation. Boden *et al.*, Endocrinology, 138:2920 (1997).

In addition to extracellular signals, such as the BMPs, intracellular  
10 signals or regulatory molecules may also play a role in the cascade of events leading to formation of new bone. One broad class of intracellular regulatory molecules are the LIM proteins, which are so named because they possess a characteristic structural motif known as the LIM domain. The LIM domain is a cysteine-rich structural motif composed of two special zinc fingers that are  
15 joined by a 2-amino acid spacer. Some proteins have only LIM domains, while others contain a variety of additional functional domains. LIM proteins form a diverse group, which includes transcription factors and cytoskeletal proteins. The primary role of LIM domains appears to be in mediating protein-protein interactions, through the formation of dimers with identical or different LIM  
20 domains, or by binding distinct proteins.

In LIM homeodomain proteins, that is, proteins having both LIM domains and a homeodomain sequence, the LIM domains function as negative regulatory elements. LIM homeodomain proteins are involved in the control of cell lineage determination and the regulation of differentiation, although LIM-  
25 only proteins may have similar roles. LIM-only proteins are also implicated in the control of cell proliferation since several genes encoding such proteins are associated with oncogenic chromosome translocations.

Humans and other mammalian species are prone to diseases or injuries that require the processes of bone repair and/or regeneration. For example,  
30 treatment of fractures would be improved by new treatment regimens that could stimulate the natural bone repair mechanisms, thereby reducing the time

required for the fractured bone to heal. In another example, individuals afflicted with systemic bone disorders, such as osteoporosis, would benefit from treatment regimens that would results in systemic formation of new bone.

Such treatment regimens would reduce the incidence of fractures arising from  
5 the loss of bone mass that is a characteristic of this disease.

For at least these reasons, extracellular factors, such as the BMPs, have been investigated for the purpose of using them to stimulate formation of new bone *in vivo*. Despite the early successes achieved with BMPs and other extracellular signalling molecules, their use entails a number of disadvantages.  
10 For example, relatively large doses of purified BMPs are required to enhance the production of new bone, thereby increasing the expense of such treatment methods. Furthermore, extracellular proteins are susceptible to degradation following their introduction into a host animal. In addition, because they are typically immunogenic, the possibility of stimulating an immune response to the  
15 administered proteins is ever present.

Due to such concerns, it would be desirable to have available treatment regimens that use an intracellular signalling molecule to induce new bone formation. Advances in the field of gene therapy now make it possible to introduce into osteogenic precursor cells, that is, cells involved in bone  
20 formation, or peripheral blood leukocytes, nucleotide fragments encoding intracellular signals that form part of the bone formation process. Gene therapy for bone formation offers a number of potential advantages: (1) lower production costs; (2) greater efficacy, compared to extracellular treatment regiments, due to the ability to achieve prolonged expression of the  
25 intracellular signal; (3) it would by-pass the possibility that treatment with extracellular signals might be hampered due to the presence of limiting numbers of receptors for those signals; (4) it permits the delivery of transfected potential osteoprogenitor cells directly to the site where localized bone formation is required; and (5) it would permit systemic bone formation, thereby  
30 providing a treatment regimen for osteoporosis and other metabolic bone diseases.

## SUMMARY OF THE INVENTION

The present invention seeks to overcome the drawbacks in the prior art by providing novel compositions and methods for inducing bone formation using an intracellular signalling molecule that participates early in the cascade of events that leads to bone formation. Applicants have discovered 10-4/RLMP (SEQ ID NO: 1, SEQ ID NO: 2), a novel LIM gene with a sequence originally isolated from stimulated rat calvarial osteoblast cultures. The gene has been cloned, sequenced and assayed for its ability to enhance the efficacy of bone mineralization *in vitro*. The protein RLMP affects mineralization of bone matrix as well as differentiation of cells into the osteoblast lineage. Unlike other known cytokines, for example, BMPs, RLMP is not a secreted protein, but is instead an intracellular signaling molecule. This feature has the advantage of providing intracellular signaling amplification as well as easier assessment of transfected cells. It is also suitable for more efficient and specific *in vivo* applications. Suitable clinical applications include enhancement of bone repair in fractures, bone defects, bone grafting, and normal homeostasis in patients presenting with osteoporosis.

Applicants have also cloned, sequenced and deduced the amino acid sequence of a corresponding human protein, named human LMP-1. The human protein demonstrates enhanced efficacy of bone mineralization *in vitro* and *in vivo*.

In addition, the applicants have characterized a truncated (short) version of LMP-1, termed HLMP-1s. This short version resulted from a point mutation in one source of a cDNA clone, providing a stop codon which truncated the protein. The short version (LMP-1s) is fully functional when expressed in cell culture and *in vivo*.

Using PCR analysis of human heart cDNA library, Applicants have identified two alternative splice variants (referred to as HLMP-2 and HLMP-3) that differ from HLMP-1 in a region between base pairs 325 and 444 in the

nucleotide sequence encoding HLMP-1. The HLMP-2 sequence has a 119 base pair deletion and an insertion of 17 base pairs in this region. Compared to HLMP-1, the nucleotide sequence encoding HLMP-3 has no deletions, but it does have the same 17 base pairs as HLMP-2, which are inserted at position  
5 444 in the HLMP-1 sequence.

Additional features and advantages of the invention will be set forth in the description which follows, and in part will be apparent from the description, or may be learned by practice of the invention. The objectives and other advantages of the invention will be realized and attained by the methods and  
10 compositions of matter particularly pointed out in the written description and claims hereof.

In one broad aspect, the invention relates to an isolated nucleic acid molecule comprising a nucleic acid sequence encoding any LIM mineralization protein, wherein the nucleic acid molecule hybridizes under standard  
15 conditions to a nucleic acid molecule complementary to the full length of SEQ. ID NO: 25, and wherein the molecule hybridizes under highly stringent conditions to a nucleic acid molecule complementary to the full length of SEQ. ID NO: 26. In a specific aspect, the isolated nucleic acid molecule encodes HLMP-1, HLMP-1s, RLMP, HLMP-2, or HLMP-3. In addition, the invention is  
20 directed to vectors comprising these nucleic acid molecules, as well as host cells comprising the vectors. In another specific aspect, the invention relates to the proteins themselves.

In a second broad aspect, the invention relates to antibody that is specific for LIM mineralization protein, including HLMP-1, HLMP-1s, RLMP,  
25 HLMP-2, and HLMP-3. In one specific aspect, the antibody is a polyclonal antibody. In another specific aspect, the antibody is a monoclonal antibody.

In a third broad aspect, the invention relates to method of inducing bone formation wherein osteogenic precursor cells are transfected with an isolated nucleic acid molecule comprising a nucleotide sequence encoding LIM  
30 mineralization protein. In one specific aspect, the isolated nucleic acid molecule is in a vector, which may be a plasmid or a virus, such as adenovirus

or retrovirus. The transfection may occur *ex vivo* or *in vivo* by direct injection of the isolated nucleic acid molecule. The transfected isolated nucleic acid molecule may encode HLMP-1, HLMP-1s, RLMP, HLMP-2, or HLMP-3.

In a further aspect, the invention relates to methods of fusing a spine by  
 5 transfecting osteogenic precursor cells with an isolated nucleic acid molecule having a nucleotide sequence encoding LIM mineralization protein, admixing the transfected osteogenic precursor cells with a matrix and contacting the matrix with the spine.

In yet another aspect, the invention relates to methods for inducing  
 10 systemic bone formation by stable transfection of host cells with the vectors of the invention.

It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are intended to provide further explanation of the invention as claimed.

15

#### ABBREVIATIONS AND DEFINITIONS

	BMP	Bone Morphogenetic Protein
	HLMP-1	Human LMP-1, also
20		designated as Human LIM
		Protein or HLMP
	HLMP-1s	Human LMP-1 Short
		(truncated) protein
	HLMPU	Human LIM Protein Unique
25		Region
	LMP	LIM mineralization protein
	MEM	Minimal essential medium
	Trm	Triamcinolone
	-GlyP	Beta-glycerolphosphate
30	RACE	Rapid Amplification of cDNA
		Ends

	RLMP	Rat LIM mineralization protein, also designated as RLMP-1
	RLMPU	Rat LIM Protein Unique Region
5	RNAasin	RNase inhibitor
	ROB	Rat Osteoblast
	10-4	Clone containing cDNA sequence for RLMP (SEQ ID NO: 2)
10	UTR	Untranslated Region
	HLMP-2	Human LMP Splice Variant 2
	HLMP-3	Human LMP Splice Variant 3
15		

## DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel mammalian LIM proteins, herein  
 20 designated LIM mineralization proteins, or LMP. The invention relates more  
 particularly to human LMP, known as HLMP or HLMP-1, or alternative splice  
 variants of human LMP, which are known as HLMP-2 or HLMP-3. The  
 applicants have discovered that these proteins enhance bone mineralization in  
 mammalian cells grown *in vitro*. When produced in mammals, LMP also  
 25 induces bone formation *in vivo*.

*Ex vivo* transfection of bone marrow cells, osteogenic precursor cells,  
 peripheral blood leukocytes, or mesenchymal stem cells with nucleic acid that  
 encodes LMP or HLMP, followed by reimplantation of the transfected cells in  
 the donor, is suitable for treating a variety of bone-related disorders or injuries.  
 30 For example, one can use this method to: augment long bone fracture repair;  
 generate bone in segmental defects; provide a bone graft substitute for



fractures; facilitate tumor reconstruction or spine fusion; and provide a local treatment (by injection) for weak or osteoporotic bone, such as in osteoporosis of the hip, vertebrae, or wrist. Transfection with LMP or HLMP-encoding nucleic acid is also useful in: the percutaneous injection of transfected marrow  
5 cells to accelerate the repair of fractured long bones; treatment of delayed union or non-unions of long bone fractures or pseudoarthrosis of spine fusions; and for inducing new bone formation in avascular necrosis of the hip or knee.

In addition to *ex vivo*-based methods of gene therapy, transfection of a recombinant DNA vector comprising a nucleic acid sequence that encodes  
10 LMP or HLMP can be accomplished *in vivo*. When a DNA fragment that encodes LMP or HLMP is inserted into an appropriate viral vector, for example, an adenovirus vector, the viral construct can be injected directly into a body site where endochondral bone formation is desired. By using a direct, percutaneous injection to introduce the LMP or HLMP sequence stimulation of  
15 bone formation can be accomplished without the need for surgical intervention either to obtain bone marrow cells (to transfect *ex vivo*) or to reimplant them into the patient at the site where new bone is required. Alden *et al.*, Neurosurgical Focus (1998), have demonstrated the utility of a direct injection method of gene therapy using a cDNA that encodes BMP-2, which was cloned  
20 into an adenovirus vector.

It is also possible to carry out *in vivo* gene therapy by directly injecting into an appropriate body site, a naked, that is, unencapsulated, recombinant plasmid comprising a nucleic acid sequence that encodes HLMP. In this embodiment of the invention, transfection occurs when the naked plasmid DNA  
25 is taken up, or internalized, by the appropriate target cells, which have been described. As in the case of *in vivo* gene therapy using a viral construct, direct injection of naked plasmid DNA offers the advantage that little or no surgical intervention is required. Direct gene therapy, using naked plasmid DNA that encodes the endothelial cell mitogen VEGF (vascular endothelial growth  
30 factor), has been successfully demonstrated in human patients. Baumgartner *et al.*, Circulation, 97(12):1114-23 (1998).

By using an adenovirus vector to deliver LMP into osteogenic cells, transient expression of LMP is achieved. This occurs because adenovirus does not incorporate into the genome of target cells that are transfected. Transient expression of LMP, that is, expression that occurs during the lifetime  
5 of the transfected target cells, is sufficient to achieve the objects of the invention. Stable expression of LMP, however, can occur when a vector that incorporates into the genome of the target cell is used as a delivery vehicle. Retrovirus-based vectors, for example, are suitable for this purpose.

Stable expression of LMP is particularly useful for treating various  
10 systemic bone-related disorders, such as osteoporosis and osteogenesis imperfecta. For this embodiment of the invention, in addition to using a vector that integrates into the genome of the target cell to deliver an LMP-encoding nucleotide sequence into target cells, LMP expression is placed under the control of a regulatable promoter. For example, a promoter that is turned on by  
15 exposure to an exogenous inducing agent, such as tetracycline, is suitable. Using this approach, one can stimulate formation of new bone on a systemic basis by administering an effective amount of the exogenous inducing agent. Once a sufficient quantity of bone mass is achieved, administration of the exogenous inducing agent is discontinued. This process may be repeated as  
20 needed to replace bone mass lost, for example, as a consequence of osteoporosis.

Antibodies specific for HLMP are particularly suitable for use in methods for assaying the osteoinductive, that is, bone-forming, potential of patient cells. In this way one can identify patients at risk for slow or poor healing of bone  
25 repair. Also, HLMP-specific antibodies are suitable for use in marker assays to identify risk factors in bone degenerative diseases, such as, for example, osteoporosis.

Following well known and conventional methods, the genes of the present invention are prepared by ligation of nucleic acid segments that  
30 encode LMP to other nucleic acid sequences, such as cloning and/or expression vectors. Methods needed to construct and analyze these

recombinant vectors, for example, restriction endonuclease digests, cloning protocols, mutagenesis, organic synthesis of oligonucleotides and DNA sequencing, have been described. For DNA sequencing DNA, the dideoxyterminator method is the preferred.

5 Many treatises on recombinant DNA methods have been published, including Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, 2nd edition (1988), Davis *et al.*, Basic Methods in Molecular Biology, Elsevier (1986), and Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience (1988). These reference manuals are  
10 specifically incorporated by reference herein.

Primer-directed amplification of DNA or cDNA is a common step in the expression of the genes of this invention. It is typically performed by the polymerase chain reaction (PCR). PCR is described in U.S. Patent No. 4,800,159 to Mullis *et al.* and other published sources. The basic principle of  
15 PCR is the exponential replication of a DNA sequence by successive cycles of primer extension. The extension products of one primer, when hybridized to another primer, becomes a template for the synthesis of another nucleic acid molecule. The primer-template complexes act as substrate for DNA polymerase, which in performing its replication function, extends the primers.  
20 The conventional enzyme for PCR applications is the thermostable DNA polymerase isolated from *Thermus aquaticus*, or Taq DNA polymerase.

Numerous variations of the basic PCR method exist, and a particular procedure of choice in any given step needed to construct the recombinant vectors of this invention is readily performed by a skilled artisan. For example,  
25 to measure cellular expression of 10-4/RLMP, RNA is extracted and reverse transcribed under standard and well known procedures. The resulting cDNA is then analyzed for the appropriate mRNA sequence by PCR.

The gene encoding the LIM mineralization protein is expressed in an expression vector in a recombinant expression system. Of course, the  
30 constructed sequence need not be the same as the original, or its complimentary sequence, but instead may be any sequence determined by the

degeneracy of the DNA code that nonetheless expresses an LMP having bone forming activity. Conservative amino acid substitutions, or other modifications, such as the occurrence of an amino-terminal methionine residue, may also be employed.

5 A ribosome binding site active in the host expression system of choice is ligated to the 5' end of the chimeric LMP coding sequence, forming a synthetic gene. The synthetic gene can be inserted into any one of a large variety of vectors for expression by ligating to an appropriately linearized plasmid. A regulatable promoter, for example, the *E. coli* lac promoter, is also suitable for  
10 the expression of the chimeric coding sequences. Other suitable regulatable promoters include trp, tac, recA, T7 and lambda promoters.

DNA encoding LMP is transfected into recipient cells by one of several standard published procedures, for example, calcium phosphate precipitation, DEAE-Dextran, electroporation or protoplast fusion, to form stable  
15 transformants. Calcium phosphate precipitation is preferred, particularly when performed as follows.

DNAs are coprecipitated with calcium phosphate according to the method of Graham and Van Der, Virology, 52:456 (1973), before transfer into cells. An aliquot of 40-50 g of DNA, with salmon sperm or calf thymus DNA as  
20 a carrier, is used for  $0.5 \times 10^6$  cells plated on a 100 mm dish. The DNA is mixed with 0.5 ml of 2X Hepes solution (280 mM NaCl, 50 mM Hepes and 1.5 mM  $\text{Na}_2\text{HPO}_4$ , pH 7.0), to which an equal volume of 2x  $\text{CaCl}_2$  (250 mM  $\text{CaCl}_2$  and 10 mM Hepes, pH 7.0) is added. A white granular precipitate, appearing after 30-40 minutes, is evenly distributed dropwise on the cells, which are allowed to  
25 incubate for 4-16 hours at 37°C. The medium is removed and the cells shocked with 15% glycerol in PBS for 3 minutes. After removing the glycerol, the cells are fed with Dulbecco's Minimal Essential Medium (DMEM) containing 10% fetal bovine serum.

DNA can also be transfected using: the DEAE-Dextran methods of  
30 Kimura *et al.*, Virology, 49:394 (1972) and Sompayrac *et al.*, Proc. Natl. Acad. Sci. USA, 78:7575 (1981); the electroporation method of Potter, Proc. Natl.

Acad. Sci. USA, 81:7161 (1984); and the protoplast fusion method of Sandri-Goddin *et al.*, Molec. Cell. Biol., 1:743 (1981).

Phosphoramidite chemistry in solid phase is the preferred method for the organic synthesis of oligodeoxynucleotides and polydeoxynucleotides. In addition, many other organic synthesis methods are available. Those methods are readily adapted by those skilled in the art to the particular sequences of the invention.

The present invention also includes nucleic acid molecules that hybridize under standard conditions to any of the nucleic acid sequences encoding the LIM mineralization proteins of the invention. "Standard hybridization conditions" will vary with the size of the probe, the background and the concentration of the nucleic acid reagents, as well as the type of hybridization, for example, *in situ*, Southern blot, or hybridization of DNA-RNA hybrids (Northern blot). The determination of "standard hybridization conditions" is within the level of skill in the art. For example, see U.S. Patent 5,580,775 to Fremeau *et al.*, herein incorporated by reference for this purpose. See also, Southern, E. M., J. Mol. Biol., 98:503 (1975), Alwine *et al.*, Meth. Enzymol., 68:220 (1979), and Sambrook *et al.*, Molecular Cloning: A laboratory Manual, 2nd edition, pp. 7.19-7.50, Cold Spring Harbor Press (1989).

One preferred set of standard hybridization conditions involves a blot that is prehybridized at 42°C for 2 hours in 50% formamide, 5X SSPE (150 mM NaCl, 10 mM Na H<sub>2</sub>PO<sub>4</sub> [pH 7.4], 1 mM EDTA [pH 8.0]), 5X Denhardt's solution (20 mg Ficoll, 20 mg polyvinylpyrrolidone and 20 mg BSA per 100 ml water), 10% dextran sulphate, 1% SDS and 100 g/ml salmon sperm DNA. A <sup>32</sup>P-labelled cDNA probe is added, and hybridization is continued for 14 hours. Afterward, the blot is washed twice with 2X SSPE, 0.1% SDS for 20 minutes at 22°C, followed by a 1 hour wash at 65°C in 0.1X SSPE, 0.1 %SDS. The blot is then dried and exposed to x-ray film for 5 days in the presence of an intensifying screen.

Under "highly stringent conditions," a probe will hybridize to its target sequence if those two sequences are substantially identical. As in the case of standard hybridization conditions, one of skill in the art can, given the level of skill in the art and the nature of the particular experiment, determine the conditions under which only substantially identical sequences will hybridize.

Another aspect of the invention includes the proteins encoded by the nucleic acid sequences. In still another embodiment, the invention relates to the identification of such proteins based on anti-LMP antibodies. In this embodiment, protein samples are prepared for Western blot analysis by lysing cells and separating the proteins by SDS-PAGE. The proteins are transferred to nitrocellulose by electroblotting as described by Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley and Sons (1987). After blocking the filter with instant nonfat dry milk (1 gm in 100 ml PBS), anti-LMP antibody is added to the filter and incubated for 1 hour at room temperature. The filter is washed thoroughly with phosphate buffered saline (PBS) and incubated with horseradish peroxidase (HRPO)-antibody conjugate for 1 hour at room temperature. The filter is again washed thoroughly with PBS and the antigen bands are identified by adding diaminobenzidine (DAB).

Monospecific antibodies are the reagent of choice in the present invention, and are specifically used to analyze patient cells for specific characteristics associated with the expression of LMP. "Monospecific antibody" as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for LMP. "Homogeneous binding" as used herein refers to the ability of the antibody species to bind to a specific antigen or epitope, such as those associated with LMP, as described above. Monospecific antibodies to LMP are purified from mammalian antisera containing antibodies reactive against LMP or are prepared as monoclonal antibodies reactive with LMP using the technique of Kohler and Milstein, Nature, 256:495-97 (1975). The LMP specific antibodies are raised by immunizing animals such as, for example, mice, rats, guinea

pigs, rabbits, goats or horses, with an appropriate concentration of LMP either with or without an immune adjuvant.

In this process, preimmune serum is collected prior to the first immunization. Each animal receives between about 0.1 mg and about 1000  
5 mg of LMP associated with an acceptable immune adjuvant, if desired. Such acceptable adjuvants include, but are not limited to, Freund's complete, Freund's incomplete, alum-precipitate, water in oil emulsion containing *Corynebacterium parvum* and tRNA adjuvants. The initial immunization consists of LMP in, preferably, Freund's complete adjuvant injected at multiple  
10 sites either subcutaneously (SC), intraperitoneally (IP) or both. Each animal is bled at regular intervals, preferably weekly, to determine antibody titer. The animals may or may not receive booster injections following the initial immunization. Those animals receiving booster injections are generally given an equal amount of the antigen in Freund's incomplete adjuvant by the same  
15 route. Booster injections are given at about three week intervals until maximal titers are obtained. At about 7 days after each booster immunization or about weekly after a single immunization, the animals are bled, the serum collected, and aliquots are stored at about -20° C.

Monoclonal antibodies (mAb) reactive with LMP are prepared by  
20 immunizing inbred mice, preferably Balb/c mice, with LMP. The mice are immunized by the IP or SC route with about 0.1 mg to about 10 mg, preferably about 1 mg, of LMP in about 0.5 ml buffer or saline incorporated in an equal volume of an acceptable adjuvant, as discussed above. Freund's complete adjuvant is preferred. The mice receive an initial immunization on day 0 and  
25 are rested for about 3-30 weeks. Immunized mice are given one or more booster immunizations of about 0.1 to about 10 mg of LMP in a buffer solution such as phosphate buffered saline by the intravenous (IV) route. Lymphocytes from antibody-positive mice, preferably splenic lymphocytes, are obtained by removing the spleens from immunized mice by standard procedures known in  
30 the art. Hybridoma cells are produced by mixing the splenic lymphocytes with an appropriate fusion partner, preferably myeloma cells, under conditions

which will allow the formation of stable hybridomas. Fusion partners may include, but are not limited to: mouse myelomas P3/NS1/Ag 4-1; MPC-11; S-194 and Sp 2/0, with Sp 2/0 being preferred. The antibody producing cells and myeloma cells are fused in polyethylene glycol, about 1000 mol. wt., at concentrations from about 30% to about 50%. Fused hybridoma cells are selected by growth in hypoxanthine, thymidine and aminopterin in supplemented Dulbecco's Modified Eagles Medium (DMEM) by procedures known in the art. Supernatant fluids are collected from growth positive wells on about days 14, 18, and 21, and are screened for antibody production by an immunoassay such as solid phase immunoradioassay (SPIRA) using LMP as the antigen. The culture fluids are also tested in the Ouchterlony precipitation assay to determine the isotype of the mAb. Hybridoma cells from antibody positive wells are cloned by a technique such as the soft agar technique of MacPherson, "Soft Agar Techniques", in Tissue Culture Methods and Applications, Kruse and Paterson (eds.), Academic Press (1973). See, also, Harlow *et al.*, Antibodies: A Laboratory Manual, Cold Spring Laboratory (1988).

Monoclonal antibodies may also be produced *in vivo* by injection of pristane- primed Balb/c mice, approximately 0.5 ml per mouse, with about  $2 \times 10^6$  to about  $6 \times 10^6$  hybridoma cells about 4 days after priming. Ascites fluid is collected at approximately 8-12 days after cell transfer and the monoclonal antibodies are purified by techniques known in the art.

*In vitro* production in anti-LMP mAb is carried out by growing the hybridoma cell line in DMEM containing about 2% fetal calf serum to obtain sufficient quantities of the specific mAb. The mAb are purified by techniques known in the art.

Antibody titers of ascites or hybridoma culture fluids are determined by various serological or immunological assays, which include, but are not limited to, precipitation, passive agglutination, enzyme-linked immunosorbent antibody (ELISA) technique and radioimmunoassay (RIA) techniques. Similar assays



are used to detect the presence of the LMP in body fluids or tissue and cell extracts.

It is readily apparent to those skilled in the art that the above described methods for producing monospecific antibodies may be utilized to produce  
5 antibodies specific for polypeptide fragments of LMP, full-length nascent LMP polypeptide, or variants or alleles thereof.

In another embodiment, the invention is directed to alternative splice variants of HLMP-1. PCR analysis of human heart cDNA revealed mRNA for two HLMP alternative splice variants, named HLMP-2 and HLMP-3, that differ  
10 from HLMP-1 in a region between base pairs 325 and 444 in the hLMP-1 sequence. The HLMP-2 sequence has a 119 base pair deletion and an insertion of 17 base pairs in this region. These changes preserve the reading frame, resulting in a 423 amino acid protein, which compared to HLMP-1, has a net loss of 34 amino acids (40 amino acids deleted plus 6 inserted amino  
15 acids). HLMP-2 contains the c-terminal LIM domains that are present in HLMP-1.

Compared to HLMP-1, HLMP-3 has no deletions, but it does have the same 17 base pair insertion at position 444. This insertion shifts the reading frame, causing a stop codon at base pairs 459-461. As a result, HLMP-3  
20 encodes a protein of 153 amino acids. This protein lacks the c-terminal LIM domains that are present in HLMP-1 and HLMP-2. The predicted size of the proteins encoded by HLMP-2 and HLMP-3 was confirmed by western blot analysis.

PCR analysis of the tissue distribution of the three splice variants  
25 revealed that they are differentially expressed, with specific isoforms predominating in different tissues. HLMP-1 is apparently the predominant form expressed in leukocytes, spleen, lung, placenta, and fetal liver. HLMP-2 appears to be the predominant isoform in skeletal muscle, bone marrow, and heart tissue. HLMP-3, however, was not the predominant isoform in any tissue  
30 examined.

Overexpression of HLMP-3 in secondary rat osteoblast cultures induced bone nodule formation ( $287 \pm 56$ ) similar to the effect seen for glucicorticoid ( $272 \pm 7$ ) and HLMP-1 ( $232 \pm 200$ ). Since HLMP-3 lacks the C-terminal LIM domains, these regions are not required for osteoinductive activity.

5 Overexpression of HLMP-2, however, did not induce nodule formation ( $11 \pm 3$ ). These data suggest that the amino acids encoded by the deleted 119 base pairs are necessary for osteoinduction. The data also suggest that the distribution of HLMP splice variants may be important for tissue-specific function. Surprisingly, we have shown that HLMP-2 inhibits steroid-induced  
10 osteoblast formation in secondary rat osteoblast cultures. Therefore, HLMP-2 will have therapeutic utility in clinical situations where bone formation is not desirable.

On July 22, 1997, a sample of 10-4/RLMP in a vector designated pCMV2/RLMP (which is vector pRc/CMV2 with insert 10-4 clone/RLMP) was  
15 deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD 20852. The culture accession number for that deposit is 209153. On March 19, 1998, a sample of the vector pHis-A with insert HLPM-1s was deposited at the American Type Culture Collection ("ATCC"). The culture accession number for that deposit is 209698. On April 14, 2000,  
20 samples of plasmids pHAhLMP-2 (vector pHisA with cDNA insert derived from human heart muscle cDNA with HLMP-2) and pHAhLMP-3 (vector pHisA with cDNA insert derived from human heart muscle cDNA with HLMP-3) were deposited with the ATCC, 10801 University Blvd., Manassas, VA, 20110-2209, USA, under the conditions of the Budapest treaty. The accession numbers for  
25 these deposits are \_\_\_\_\_ and \_\_\_\_\_, respectively. These deposits, as required by the Budapest Treaty, will be maintained in the ATCC for at least 30 years and will be made available to the public upon the grant of a patent disclosing them. It should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent  
30 rights granted by government action.

In assessing the nucleic acids, proteins, or antibodies of the invention, enzyme assays, protein purification, and other conventional biochemical methods are employed. DNA and RNA are analyzed by Southern blotting and Northern blotting techniques, respectively. Typically, the samples analyzed are size fractionated by gel electrophoresis. The DNA or RNA in the gels are then transferred to nitrocellulose or nylon membranes. The blots, which are replicas of sample patterns in the gels, were then hybridized with probes. Typically, the probes are radiolabelled, preferably with  $^{32}\text{P}$ , although one could label the probes with other signal-generating molecules known to those in the art. Specific bands of interest can then be visualized by detection systems, such as autoradiography.

For purposes of illustrating preferred embodiments of the present invention, the following, non-limiting examples are included. These results demonstrate the feasibility of inducing or enhancing the formation of bone using the LIM mineralization proteins of the invention, and the isolated nucleic acid molecules encoding those proteins.

#### Example 1: Calvarial Cell Culture

Rat calvarial cells, also known as rat osteoblasts ("ROB"), were obtained from 20-day pre-parturition rats as previously described. Boden *et al.*, Endocrinology, 137(8):3401-07 (1996). Primary cultures were grown to confluence (7 days), trypsinized, and passed into 6-well plates ( $1 \times 10^5$  cells/35 mm well) as first subculture cells. The subculture cells, which were confluent at day 0, were grown for an additional 7 days. Beginning on day 0, media were changed and treatments (Tm and/or BMPs) were applied, under a laminar flow hood, every 3 or 4 days. The standard culture protocol was as follows: days 1-7, MEM, 10% FBS, 50 g/ml ascorbic acid,  $\pm$  stimulus; days 8-14, BGJb medium, 10% FBS, 5mM -GlyP (as a source of inorganic phosphate to permit mineralization). Endpoint analysis of bone nodule formation and osteocalcin secretion was performed at day 14. The dose of BMP was chosen as 50 ng/ml

based on pilot experiments in this system that demonstrated a mid-range effect on the dose-response curve for all BMPs studied.

#### EXAMPLE 2: Antisense Treatment and Cell Culture

5

To explore the potential functional role of LMP-1 during membranous bone formation, we synthesized an antisense oligonucleotide to block LMP-1 mRNA translation and treated secondary osteoblast cultures that were undergoing differentiation initiated by glucocorticoid. Inhibition of RLMP  
10 expression was accomplished with a highly specific antisense oligonucleotide (having no significant homologies to known rat sequences) corresponding to a 25 bp sequence spanning the putative translational start site (SEQ ID NO: 42). Control cultures either did not receive oligonucleotide or they received sense oligonucleotide. Experiments were performed in the presence (preincubation)  
15 and absence of lipofectamine. Briefly, 22 g of sense or antisense RLMP oligonucleotide was incubated in MEM for 45 minutes at room temperature. Following that incubation, either more MEM or pre-incubated lipofectamine/MEM (7% v/v; incubated 45 minutes at room temperature) was added to achieve an oligonucleotide concentration of 0.2 M. The resulting  
20 mixture was incubated for 15 minutes at room temperature. Oligonucleotide mixtures were then mixed with the appropriate medium, that is, MEM/Ascorbate/ $\pm$ Trm, to achieve a final oligonucleotide concentration of 0.1 M.

Cells were incubated with the appropriate medium ( $\pm$ stimulus) in the  
25 presence or absence of the appropriate oligonucleotides. Cultures originally incubated with lipofectamine were re-fed after 4 hours of incubation (37°C; 5% CO<sub>2</sub>) with media containing neither lipofectamine nor oligonucleotide. All cultures, especially cultures receiving oligonucleotide, were re-fed every 24 hours to maintain oligonucleotide levels.

30 LMP-1 antisense oligonucleotide inhibited mineralized nodule formation and osteocalcin secretion in a dose-dependent manner, similar to the effect of

BMP-6 oligonucleotide. The LMP-1 antisense block in osteoblast differentiation could not be rescued by addition of exogenous BMP-6, while the BMP-6 antisense oligonucleotide inhibition was reversed with addition of BMP-6. This experiment further confirmed the upstream position of LMP-1 relative to BMP-6 in the osteoblast differentiation pathway. LMP-1 antisense oligonucleotide also inhibited spontaneous osteoblast differentiation in primary rat osteoblast cultures.

### EXAMPLE 3: Quantitation of Mineralized Bone Nodule Formation

10

Cultures of ROBs prepared according to Examples 1 and 2 were fixed overnight in 70% ethanol and stained with von Kossa silver stain. A semi-automated computerized video image analysis system was used to quantitate nodule count and nodule area in each well. Boden *et al.*, Endocrinology, 137(8):3401-07 (1996). These values were then divided to calculate the area per nodule values. This automated process was validated against a manual counting technique and demonstrated a correlation coefficient of 0.92 ( $p < 0.000001$ ). All data are expressed as the mean  $\pm$  standard error of the mean (S.E.M.) calculated from 5 or 6 wells at each condition. Each experiment was confirmed at least twice using cells from different calvarial preparations.

### EXAMPLE 4: Quantitation of Osteocalcin Secretion

Osteocalcin levels in the culture media were measured using a competitive radioimmunoassay with a monospecific polyclonal antibody (Pab) raised in our laboratory against the C-terminal nonapeptide of rat osteocalcin as described in Nanes *et al.*, Endocrinology, 127:588 (1990). Briefly, 1 g of nonapeptide was iodinated with 1 mCi  $^{125}\text{I}$ -Na by the lactoperoxidase method. Tubes containing 200  $\mu\text{l}$  of assay buffer (0.02 M sodium phosphate, 1 mM EDTA, 0.001% thimerosal, 0.025% BSA) received media taken from cell cultures or osteocalcin standards (0 - 12,000 fmole) at 100  $\mu\text{l}$ /tube in assay

buffer. The Pab (1:40,000; 100 l) was then added, followed by the iodinated peptide (12,000 cpm; 100 l). Samples tested for non-specific binding were prepared similarly but contained no antibody.

Bound and free PABs were separated by the addition of 700 l goat anti-  
5 rabbit IgG, followed by incubation for 18 hours at 4°C. After samples were centrifuged at 1200 rpm for 45 minutes, the supernatants were decanted and the precipitates counted in a gamma counter. Osteocalcin values were reported in fmole/100 l, which was then converted to pmole/ml medium (3-day production) by dividing those values by 100. Values were expressed as the  
10 mean  $\pm$  S.E.M. of triplicate determinations for 5-6 wells for each condition. Each experiment was confirmed at least two times using cells from different calvarial preparations.

#### EXAMPLE 5: Effect of Trm and RLMP on Mineralization *In Vitro*

15

There was little apparent effect of either the sense or antisense oligonucleotides on the overall production of bone nodules in the non-stimulated cell culture system. When ROB's were stimulated with Trm, however, the antisense oligonucleotide to RLMP inhibited mineralization of  
20 nodules by > 95%. The addition of exogenous BMP-6 to the oligonucleotide-treated cultures did not rescue the mineralization of RLMP-antisense-treated nodules.

Osteocalcin has long been synonymous with bone mineralization, and osteocalcin levels have been correlated with nodule production and  
25 mineralization. The RLMP-antisense oligonucleotide significantly decreases osteocalcin production, but the nodule count in antisense-treated cultures does not change significantly. In this case, the addition of exogenous BMP-6 only rescued the production of osteocalcin in RLMP-antisense-treated cultures by 10-15%. This suggests that the action of RLMP is downstream of, and more  
30 specific than, BMP-6.

EXAMPLE 6: Harvest and Purification of RNA

Cellular RNA from duplicate wells of ROBs (prepared according to  
5 Examples 1 and 2 in 6-well culture dishes) was harvested using 4M guanidine  
isothiocyanate (GIT) solution to yield statistical triplicates. Briefly, culture  
supernatant was aspirated from the wells, which were then overlaid with 0.6  
ml of GIT solution per duplicate well harvest. After adding the GIT solution, the  
plates were swirled for 5-10 seconds (being as consistent as possible).  
10 Samples were saved at -70°C for up to 7 days before further processing.

RNA was purified by a slight modification of standard methods  
according to Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd  
Ed., chapter 7.19, Cold Spring Harbor Press (1989). Briefly, thawed samples  
received 60 l 2.0 M sodium acetate (pH 4.0), 550 l phenol (water saturated)  
15 and 150 l chloroform:isoamyl alcohol (49:1). After vortexing, the samples were  
centrifuged (10000 x g; 20 minutes; 4°C), the aqueous phase transferred to a  
fresh tube, 600 l isopropanol was added and the RNA precipitated overnight at  
-20°C.

Following the overnight incubation, the samples were centrifuged  
20 (10000 x g; 20 minutes) and the supernatant was aspirated gently. The pellets  
were resuspended in 400 l DEPC-treated water, extracted once with  
phenol:chloroform (1:1), extracted with chloroform:isoamyl alcohol (24:1) and  
precipitated overnight at -20°C after addition of 40 l sodium acetate (3.0 M; pH  
5.2) and 1.0 ml absolute ethanol. To recover the cellular RNA, the samples  
25 were centrifuged (10000 x g; 20 min), washed once with 70% ethanol, air dried  
for 5-10 minutes and resuspended in 20 l of DEPC-treated water. RNA  
concentrations were calculated from optical densities that were determined  
with a spectrophotometer.

EXAMPLE 7: Reverse Transcription-Polymerase Chain Reaction

Heated total RNA (5 g in 10.5 l total volume DEPC-H<sub>2</sub>O at 65°C for 5 minutes) was added to tubes containing 4 l 5X MMLV-RT buffer, 2 l dNTPs, 2 l  
5 dT17 primer (10 pmol/ml), 0.5 l RNAsin (40U/ml) and 1 l MMLV-RT (200 units/l). The samples were incubated at 37°C for 1 hour, then at 95°C for 5 minutes to inactivate the MMLV-RT. The samples were diluted by addition of 80 l of water.

Reverse-transcribed samples (5 l) were subjected to polymerase-chain  
10 reaction using standard methodologies (50 l total volume). Briefly, samples were added to tubes containing water and appropriate amounts of PCR buffer, 25 mM MgCl<sub>2</sub>, dNTPs, forward and reverse primers for glyceraldehyde 3-phosphate dehydrogenase (GAP, a housekeeping gene) and/or BMP-6), <sup>32</sup>P-dCTP, and Taq polymerase. Unless otherwise noted, primers were  
15 standardized to run consistently at 22 cycles (94°C, 30"; 58°C, 30"; 72°C, 20").

EXAMPLE 8: Quantitation of RT-PCR Products by Polyacrylamide Gel Electrophoresis (PAGE) and PhosphorImager Analysis

20

RT-PCR products received 5 l/tube loading dye, were mixed, heated at 65°C for 10 min and centrifuged. Ten l of each reaction was subjected to PAGE (12% polyacrylamide:bis; 15 V/well; constant current) under standard conditions. Gels were then incubated in gel preserving buffer (10% v/v  
25 glycerol, 7% v/v acetic acid, 40% v/v methanol, 43% deionized water) for 30 minutes, dried (80°C) *in vacuo* for 1-2 hours and developed with an electronically-enhanced phosphorescence imaging system for 6-24 hours. Visualized bands were analyzed. Counts per band were plotted graphically.

30



EXAMPLE 9: Differential Display PCR

RNA was extracted from cells stimulated with glucocorticoid (Trm, 1 nM). Heated, DNase-treated total RNA (5  $\mu$ g in 10.5  $\mu$ l total volume in DEPC-H<sub>2</sub>O at 65°C for 5 minutes) was reverse transcribed as described in Example 7, but H-T<sub>11</sub>M (SEQ ID. NO: 4) was used as the MMLV-RT primer. The resulting cDNAs were PCR-amplified as described above, but with various commercial primer sets (for example, H-T<sub>11</sub>G (SEQ ID NO: 4) and H-AP-10 (SEQ ID. NO: 5); GenHunter Corp, Nashville, TN). Radiolabelled PCR products were fractionated by gel electrophoresis on a DNA sequencing gel. After electrophoresis, the resulting gels were dried *in vacuo* and autoradiographs were exposed overnight. Bands representing differentially-expressed cDNAs were excised from the gel and reamplified by PCR using the method of Conner *et al.*, Proc. Natl. Acad. Sci. USA, 88:278 (1983). The products of PCR reamplification were cloned into the vector PCR-II (TA cloning kit; InVitrogen, Carlsbad, CA).

EXAMPLE 10: Screening of a UMR 106 Rat Osteosarcoma Cell cDNA Library

A UMR 106 library (2.5 x 10<sup>10</sup> pfu/ml) was plated at 5 x 10<sup>4</sup> pfu/ml onto agar plates (LB bottom agar) and the plates were incubated overnight at 37°C. Filter membranes were overlaid onto plates for two minutes. Once removed, the filters were denatured, rinsed, dried and UV cross-linked. The filters were then incubated in pre-hybridization buffer (2X PIPES [pH 6.5], 5% formamide, 1% SDS and 100  $\mu$ g/ml denatured salmon sperm DNA) for 2 h at 42°C. A 260 base-pair radiolabelled probe (SEQ ID NO: 3; <sup>32</sup>P labelled by random priming) was added to the entire hybridization mix/filters, followed by hybridization for 18 hours at 42°C. The membranes were washed once at room temperature (10 min, 1 x SSC, 0.1% SDS) and three times at 55°C (15 min, 0.1 x SSC, 0.1% SDS).

After they were washed, the membranes were analyzed by autoradiography as described above. Positive clones were plaque purified. The procedure was repeated with a second filter for four minutes to minimize spurious positives. Plaque-purified clones were rescued as lambda SK(-) phagemids. Cloned cDNAs were sequenced as described below.

#### EXAMPLE 11: Sequencing of Clones

Cloned cDNA inserts were sequenced by standard methods. Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience (1988). Briefly, appropriate concentrations of termination mixture, template and reaction mixture were subjected to an appropriate cycling protocol (95°C,30s; 68°C,30s; 72°C,60s; x 25). Stop mixture was added to terminate the sequencing reactions. After heating at 92°C for 3 minutes, the samples were loaded onto a denaturing 6% polyacrylamide sequencing gel (29:1 acrylamide:bis-acrylamide). Samples were electrophoresed for about 4 hours at 60 volts, constant current. After electrophoresis, the gels were dried *in vacuo* and autoradiographed.

The autoradiographs were analyzed manually. The resulting sequences were screened against the databases maintained by the National Center for Biotechnology Information (NIH, Bethesda, MD; <http://www.ncbi.nlm.nih.gov/>) using the BLASTn program set with default parameters. Based on the sequence data, new sequencing primers were prepared and the process was repeated until the entire gene had been sequenced. All sequences were confirmed a minimum of three times in both orientations.

Nucleotide and amino acid sequences were also analyzed using the PCGENE software package (version 16.0). Per cent homology values for nucleotide sequences were calculated by the program NALIGN, using the following parameters: weight of non-matching nucleotides, 10; weight of non-matching gaps, 10; maximum number of nucleotides considered, 50; and minimum number of nucleotides considered, 50.

For amino acid sequences, per cent homology values were calculated using PALIGN. A value of 10 was selected for both the open gap cost and the unit gap cost.

5    EXAMPLE 12: Cloning of RLMP cDNA

The differential display PCR amplification products described in Example 9 contained a major band of approximately 260 base pairs. This sequence was used to screen a rat osteosarcoma (UMR 106) cDNA library.

- 10   Positive clones were subjected to nested primer analysis to obtain the primer sequences necessary for amplifying the full length cDNA. (SEQ. ID NOs: 11, 12, 29, 30 and 31) One of those positive clones selected for further study was designated clone 10-4.

- Sequence analysis of the full-length cDNA in clone 10-4, determined by  
15   nested primer analysis, showed that clone 10-4 contained the original 260 base-pair fragment identified by differential display PCR. Clone 10-4 (1696 base pairs; SEQ ID NO: 2) contains an open reading frame of 1371 base pairs encoding a protein having 457 amino acids (SEQ ID NO: 1). The termination codon, TGA, occurs at nucleotides 1444-1446. The polyadenylation signal at  
20   nucleotides 1675-1680, and adjacent poly(A)<sup>+</sup> tail, was present in the 3' noncoding region. There were two potential N-glycosylation sites, Asn-Lys-Thr and Asn-Arg-Thr, at amino acid positions 113-116 and 257-259 in SEQ ID NO: 1, respectively. Two potential cAMP- and cGMP-dependent protein kinase phosphorylation sites, Ser and Thr, were found at amino acid positions 191  
25   and 349, respectively. There were five potential protein kinase C phosphorylation sites, Ser or Thr, at amino acid positions 3, 115, 166, 219, 442. One potential ATP/GTP binding site motif A (P-loop), Gly-Gly-Ser-Asn-Asn-Gly-Lys-Thr, was determined at amino acid positions 272-279.

- In addition, two highly conserved putative LIM domains were found at  
30   amino acid positions 341-391 and 400-451. The putative LIM domains in this newly identified rat cDNA clone showed considerable homology with the LIM

domains of other known LIM proteins. However, the overall homology with other rat LIM proteins was less than 25%. RLMP (also designated 10-4) has 78.5% amino acid homology to the human enigma protein (see U.S. Patent No. 5,504,192), but only 24.5% and 22.7% amino acid homology to its closest rat homologs, CLP-36 and RIT-18, respectively.

#### EXAMPLE 13: Northern Blot Analysis of RLMP Expression

Thirty  $\mu$ g of total RNA from ROB cells, prepared according to Examples 1 and 2, was size fractionated by formaldehyde gel electrophoresis in 1% agarose flatbed gels and osmotically transblotted to nylon membranes. The blot was probed with a 600 base pair EcoR1 fragment of full-length 10-4 cDNA labeled with  $^{32}$ P-dCTP by random priming.

Northern blot analysis showed a 1.7 kb mRNA species that hybridized with the RLMP probe. RLMP mRNA was up-regulated approximately 3.7-fold in ROB cells after 24 hours exposure to BMP-6. No up-regulation of RLMP expression was seen in BMP-2 or BMP-4-stimulated ROB cells at 24 hours.

#### EXAMPLE 14: Statistical Methods

For each reported nodule/osteocalcin result, data from 5-6 wells from a representative experiment were used to calculate the mean  $\pm$  S.E.M. Graphs may be shown with data normalized to the maximum value for each parameter to allow simultaneous graphing of nodule counts, mineralized areas and osteocalcin.

For each reported RT-PCR, RNase protection assay or Western blot analysis, data from triplicate samples of representative experiments, were used to determine the mean  $\pm$  S.E.M. Graphs may be shown normalized to either day 0 or negative controls and expressed as fold-increase above control values.

- Statistical significance was evaluated using a one-way analysis of variance with post-hoc multiple comparison corrections of Bonferroni as appropriate. D. V. Huntsberger, "The Analysis of Variance," in Elements of Statistical Variance, P. Billingsley (ed.), pp. 298-330, Allyn & Bacon Inc., Boston, MA (1977) and Sigmastat, Jandel Scientific, Corte Madera, CA. Alpha levels for significance were defined as  $p < 0.05$ .

EXAMPLE 15: Detection of Rat LIM Mineralization Protein by Western Blot Analysis

10

Polyclonal antibodies were prepared according to the methods of England *et al.*, Biochim.Biophys. Acta, 623:171 (1980) and Timmer *et al.*, J. Biol. Chem., 268:24863 (1993).

- HeLa cells were transfected with pCMV2/RLMP. Protein was harvested from the transfected cells according to the method of Hair *et al.*, Leukemia Research, 20:1 (1996). Western Blot Analysis of native RLMP was performed as described by Towbin *et al.*, Proc. Natl. Acad. Sci. USA, 76:4350 (1979).

EXAMPLE 16: Synthesis of the Rat LMP-Unique (RLMPU) derived Human PCR product

20

- Based on the sequence of the rat LMP-1 cDNA, forward and reverse PCR primers (SEQ ID NOs: 15 and 16) were synthesized and a unique 223 base-pair sequence was PCR amplified from the rat LMP-1 cDNA. A similar PCR product was isolated from human MG63 osteosarcoma cell cDNA with the same PCR primers.

- RNA was harvested from MG63 osteosarcoma cells grown in T-75 flasks. Culture supernatant was removed by aspiration and the flasks were overlaid with 3.0 ml of GIT solution per duplicate, swirled for 5-10 seconds, and the resulting solution was transferred to 1.5 ml eppendorf tubes (5 tubes with 0.6 ml/tube). RNA was purified by a slight modification of standard

methods, for example, see Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, chapter 7, page 19, Cold Spring Harbor Laboratory Press (1989) and Boden *et al.*, *Endocrinology*, 138:2820-28 (1997). Briefly, the 0.6 ml samples received 60  $\mu$ l 2.0 M sodium acetate (pH 4.0), 550  $\mu$ l water saturated phenol and 5 150  $\mu$ l chloroform:isoamyl alcohol (49:1). After addition of those reagents, the samples were vortexed, centrifuged (10000 x g; 20 min; 4°C) and the aqueous phase transferred to a fresh tube. Isopropanol (600  $\mu$ l) was added and the RNA was precipitated overnight at -20°C. The samples were centrifuged (10000 x g; 20 minutes) and the supernatant was aspirated gently. The pellets were 10 resuspended in 400  $\mu$ l of DEPC-treated water, extracted once with phenol:chloroform (1:1), extracted with chloroform:isoamyl alcohol (24:1) and precipitated overnight at -20°C in 40  $\mu$ l sodium acetate (3.0 M; pH 5.2) and 1.0 ml absolute ethanol. After precipitation, the samples were centrifuged (10000 x g; 20 min), washed once with 70% ethanol, air dried for 5-10 minutes and 15 resuspended in 20  $\mu$ l of DEPC-treated water. RNA concentrations were derived from optical densities.

Total RNA (5  $\mu$ g in 10.5  $\mu$ l total volume in DEPC-H<sub>2</sub>O) was heated at 65°C for 5 minutes, and then added to tubes containing 4  $\mu$ l 5X MMLV-RT buffer, 2  $\mu$ l dNTPs, 2  $\mu$ l dT17 primer (10 pmol/ml), 0.5  $\mu$ l RNAsin (40 U/ml) and 1  $\mu$ l MMLV-RT 20 (200 units/ $\mu$ l). The reactions were incubated at 37°C for 1 hour. Afterward, the MMLV-RT was inactivated by heating at 95°C for 5 minutes. The samples were diluted by addition of 80  $\mu$ l water.

Transcribed samples (5  $\mu$ l) were subjected to polymerase-chain reaction using standard methodologies (50  $\mu$ l total volume). Boden *et al.*, *Endocrinology*, 25 138:2820-28 (1997); Ausubel *et al.*, "Quantitation of rare DNAs by the polymerase chain reaction", in *Current Protocols in Molecular Biology*, chapter 15.31-1, Wiley & Sons, Trenton, NJ (1990). Briefly, samples were added to tubes containing water and appropriate amounts of PCR buffer (25 mM MgCl<sub>2</sub>, dNTPs, forward and reverse primers (for RLMPU; SEQ ID NOs: 15 and 16), 30 <sup>32</sup>P-dCTP, and DNA polymerase. Primers were designed to run consistently at

22 cycles for radioactive band detection and 33 cycles for amplification of PCR product for use as a screening probe (94°C, 30 sec, 58°C, 30 sec; 72°C, 20 sec).

- Sequencing of the agarose gel-purified MG63 osteosarcoma-derived  
5 PCR product gave a sequence more than 95% homologous to the RLMPU PCR product. That sequence is designated HLMP unique region (HLMPU; SEQ ID NO: 6).

EXAMPLE 17: Screening of reverse-transcriptase-derived MG63 cDNA

10

- Screening was performed with PCR using specific primers (SEQ ID NOs: 16 and 17) as described in Example 7. A 717 base-pair MG63 PCR product was agarose gel purified and sequenced with the given primers (SEQ. ID NOs: 12, 15, 16, 17, 18, 27 and 28). Sequences were confirmed a  
15 minimum of two times in both directions. The MG63 sequences were aligned against each other and then against the full-length rat LMP cDNA sequence to obtain a partial human LMP cDNA sequence (SEQ ID NO: 7).

EXAMPLE 18: Screening of a Human Heart cDNA Library

20

- Based on Northern blot experiments, it was determined that LMP-1 is expressed at different levels by several different tissues, including human heart muscle. A human heart cDNA library was therefore examined. The library was plated at  $5 \times 10^4$  pfu/ml onto agar plates (LB bottom agar) and plates were  
25 grown overnight at 37° C. Filter membranes were overlaid onto the plates for two minutes. Afterward, the filters denatured, rinsed, dried, UV cross-linked and incubated in pre-hybridization buffer (2X PIPES [pH 6.5]; 5% formamide, 1% SDS, 100 g/ml denatured salmon sperm DNA) for 2 h at 42°C. A radiolabelled, LMP-unique, 223 base-pair probe ( $^{32}\text{P}$ , random primer labelling;  
30 SEQ ID NO: 6) was added and hybridized for 18 h at 42°C. Following hybridization, the membranes were washed once at room temperature (10 min,

1 x SSC, 0.1% SDS) and three times at 55°C (15 min, 0.1 x SSC, 0.1% SDS). Double-positive plaque-purified heart library clones, identified by autoradiography, were rescued as lambda phagemids according to the manufacturers' protocols (Stratagene, La Jolla, CA).

5           Restriction digests of positive clones yielded cDNA inserts of varying sizes. Inserts greater than 600 base-pairs in length were selected for initial screening by sequencing. Those inserts were sequenced by standard methods as described in Example 11.

          One clone, number 7, was also subjected to automated sequence  
10   analysis using primers corresponding to SEQ ID NOs: 11-14, 16 and 27. The sequences obtained by these methods were routinely 97-100% homologous. Clone 7 (Partial Human LMP-1 cDNA from a heart library; SEQ. ID NO: 8) contained sequence that was more than 87% homologous to the rat LMP cDNA sequence in the translated region.

15

EXAMPLE 19: Determination of Full-Length Human LMP-1 cDNA

          Overlapping regions of the MG63 human osteosarcoma cell cDNA sequence and the human heart cDNA clone 7 sequence were used to align  
20   those two sequences and derive a complete human cDNA sequence of 1644 base-pairs. NALIGN, a program in the PCGENE software package, was used to align the two sequences. The overlapping regions of the two sequences constituted approximately 360 base-pairs having complete homology except for a single nucleotide substitution at nucleotide 672 in the MG63 cDNA (SEQ ID  
25   NO: 7) with clone 7 having an "A" instead of a "G" at the corresponding nucleotide 516 (SEQ ID NO: 8).

          The two aligned sequences were joined using SEQIN, another subprogram of PCGENE, using the "G" substitution of the MG63 osteosarcoma cDNA clone. The resulting sequence is shown in SEQ ID NO: 9. Alignment of  
30   the novel human-derived sequence with the rat LMP-1 cDNA was accomplished with NALIGN. The full-length human LMP-1 cDNA sequence



(SEQ. ID NO: 9) is 87.3% homologous to the translated portion of rat LMP-1 cDNA sequence.

EXAMPLE 20: Determination of Amino Acid Sequence of Human LMP-1

5

The putative amino acid sequence of human LMP-1 was determined with the PCGENE subprogram TRANSL. The open reading frame in SEQ ID NO: 9 encodes a protein comprising 457 amino acids (SEQ. ID NO: 10). Using the PCGENE subprogram Palign, the human LMP-1 amino acid sequence was  
10 found to be 94.1% homologous to the rat LMP-1 amino acid sequence.

EXAMPLE 21: Determination of the 5 Prime Untranslated Region of the Human LMP cDNA

15 MG63 5' cDNA was amplified by nested RT-PCR of MG63 total RNA using a 5' rapid amplification of cDNA ends (5' RACE) protocol. This method included first strand cDNA synthesis using a lock-docking oligo (dT) primer with two degenerate nucleotide positions at the 3' end (Chenchik *et al.*, CLONTECHniques, X:5 (1995); Borson *et al.*, PC Methods Applic., 2:144  
20 (1993)). Second-strand synthesis is performed according to the method of Gubler *et al.*, Gene, 25:263 (1983), with a cocktail of *Escherichia coli* DNA polymerase I, RNase H, and *E. coli* DNA ligase. After creation of blunt ends with T4 DNA polymerase, double-stranded cDNA was ligated to the fragment (5' -CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCGGGCAGGT- 3')  
25 (SEQ.ID NO: 19). Prior to RACE, the adaptor-ligated cDNA was diluted to a concentration suitable for Marathon RACE reactions (1:50). Adaptor-ligated double-stranded cDNA was then ready to be specifically cloned.

First-round PCR was performed with the adaptor-specific oligonucleotide, 5'-CCATCCTAATACGACTCACTATAGGGC- 3' (AP1) (SEQ.ID  
30 NO: 20) as sense primer and a Gene Specific Primer (GSP) from the unique region described in Example 16 (HLMPU). The second round of PCR was

performed using a nested primers GSP1-HLMPU (antisense/reverse primer) (SEQ. ID NO: 23) and GSP2-HLMPUF (SEQ. ID NO: 24) (see Example 16; sense/forward primer). PCR was performed using a commercial kit (Advantage cDNA PCR core kit; CloneTech Laboratories Inc., Palo Alto, CA) that utilizes  
5 an antibody-mediated, but otherwise standard, hot-start protocol. PCR conditions for MG63 cDNA included an initial hot-start denaturation (94°C, 60 sec) followed by: 94°C, 30 sec; 60°C, 30 sec; 68°C, 4 min; 30 cycles. The first-round PCR product was approximately 750 base-pairs in length whereas the nested PCR product was approximately 230 base-pairs. The first-round PCR  
10 product was cloned into linearized pCR 2.1 vector (3.9 Kb). The inserts were sequenced in both directions using M13 Forward and Reverse primers (SEQ. ID NO: 11; SEQ. ID NO: 12)

EXAMPLE 22: Determination of Full-length Human LMP-1 cDNA with 5 Prime  
15 UTR

Overlapping MG63 human osteosarcoma cell cDNA 5'-UTR sequence (SEQ ID NO: 21), MG63 717 base-pair sequence (Example 17; SEQ ID NO: 8) and human heart cDNA clone 7 sequence (Example 18) were aligned to derive  
20 a novel human cDNA sequence of 1704 base-pairs (SEQ.ID NO: 22). The alignment was accomplished with NALIGN, (both PCGENE and Omiga 1.0; Intelligenetics). Over-lapping sequences constituted nearly the entire 717 base-pair region (Example 17) with 100% homology. Joining of the aligned sequences was accomplished with SEQIN.

25

EXAMPLE 23: Construction of LIM Protein Expression Vector

The construction of pHIS-5ATG LMP-1s expression vector was carried out with the sequences described in Examples 17 and 18. The 717 base-pair  
30 clone (Example 17; SEQ ID NO: 7) was digested with ClaI and EcoRV. A small fragment (~250 base-pairs) was gel purified. Clone 7 (Example 18; SEQ

ID NO: 8) was digested with Clal and Xbal and a 1400 base-pair fragment was gel purified. The isolated 250 base-pair and 1400 base-pair restriction fragments were ligated to form a fragment of ~1650 base-pairs.

Due to the single nucleotide substitution in Clone 7 (relative to the 717  
5 base-pair PCR sequence and the original rat sequence) a stop codon at translated base-pair 672 resulted. Because of this stop codon, a truncated (short) protein was encoded, hence the name LMP-1s. This was the construct used in the expression vector (SEQ ID NO: 32). The full length cDNA  
10 sequence with 5' UTR (SEQ ID NO: 33) was created by alignment of SEQ ID NO: 32 with the 5' RACE sequence (SEQ ID NO: 21). The amino acid sequence of LMP-1s (SEQ ID NO: 34) was then deduced as a 223 amino acid protein and confirmed by Western blot (as in Example 15) to run at the predicted molecular weight of ~ 23.7 kD.

The pHis-ATG vector (InVitrogen, Carlsbad, CA) was digested with  
15 EcoRV and Xbal. The vector was recovered and the 1650 base-pair restriction fragment was then ligated into the linearized pHis-ATG. The ligated product was cloned and amplified. The pHis-ATG-LMP-1s Expression vector, also designated pHIS-A with insert HLMP-1s, was purified by standard methods.

20 EXAMPLE 24: Induction of Bone Nodule Formation and Mineralization *In vitro* with LMP Expression Vector

Rat Calvarial cells were isolated and grown in secondary culture according to Example 1. Cultures were either unstimulated or stimulated with  
25 glucocorticoid (GC) as described in Example 1. A modification of the Superfect Reagent (Qiagen, Valencia, CA) transfection protocol was used to transfect 3 g/well of each vector into secondary rat calvarial osteoblast cultures according to Example 25.

Mineralized nodules were visualized by Von Kossa staining, as described in  
30 Example 3. Human LMP-1s gene product overexpression alone induced bone nodule formation (~203 nodules/well) *in vitro*. Levels of nodules were

approximately 50% of those induced by the GC positive control (~412 nodules/well). Other positive controls included the pHisA-LMP-Rat expression vector (~152 nodules/well) and the pCMV2/LMP-Rat-Fwd Expression vector (~206 nodules/well), whereas the negative controls included the pCMV2/LMP-Rat-Rev. Expression vector (~2 nodules/well) and untreated (NT) plates (~4 nodules/well). These data demonstrate that the human cDNA was at least as osteoinductive as the rat cDNA. The effect was less than that observed with GC stimulation, most likely due to suboptimal doses of Expression vector.

10

#### EXAMPLE 25: LMP-Induced Cell Differentiation *In Vitro* and *In Vivo*

The rat LMP cDNA in clone 10-4 (see Example 12) was excised from the vector by double-digesting the clone with NotI and ApaI overnight at 37°C. Vector pCMV2 MCS (InVitrogen, Carlsbad, CA) was digested with the same restriction enzymes. Both the linear cDNA fragment from clone 10-4 and pCMV2 were gel purified, extracted and ligated with T4 ligase. The ligated DNA was gel purified, extracted and used to transform *E. coli* JM109 cells for amplification. Positive agar colonies were picked, digested with NotI and ApaI and the restriction digests were examined by gel electrophoresis. Stock cultures were prepared of positive clones.

A reverse vector was prepared in analogous fashion except that the restriction enzymes used were XbaI and HindIII. Because these restriction enzymes were used, the LMP cDNA fragment from clone 10-4 was inserted into pRc/CMV2 in the reverse (that is, non-translatable) orientation. The recombinant vector produced is designated pCMV2/RLMP.

An appropriate volume of pCMV10-4 (60 nM final concentration is optimal [3 g]; for this experiment a range of 0-600 nM/well [0-30 g/well] final concentration is preferred) was resuspended in Minimal Eagle Media (MEM) to 450 l final volume and vortexed for 10 seconds. Superfect was added (7.5 l/ml final solution), the solution was vortexed for 10 seconds and then incubated at

room temperature for 10 minutes. Following this incubation, MEM supplemented with 10% FBS (1 ml/well; 6 ml/plate) was added and mixed by pipetting.

5 The resulting solution was then promptly pipetted (1 ml/well) onto washed ROB cultures. The cultures were incubated for 2 hours at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. Afterward, the cells were gently washed once with sterile PBS and the appropriate normal incubation medium was added.

Results demonstrated significant bone nodule formation in all rat cell  
10 cultures which were induced with pCMV10-4. For example, pCMV10-4 transfected cells produced 429 nodules/well. Positive control cultures, which were exposed to Trm, produced 460 nodules/well. In contrast, negative controls, which received no treatment, produced 1 nodule/well. Similarly, when  
15 cultures were transfected with pCMV10-4 (reverse), no nodules were observed.

For demonstrating *de novo* bone formation *in vivo*, marrow was aspirated from the hindlimbs of 4-5 week old normal rats (rnu/+; heterozygous for recessive athymic condition). The aspirated marrow cells were washed in alpha MEM, centrifuged, and RBCs were lysed by resuspending the pellet in  
20 0.83% NH<sub>4</sub>Cl in 10 mM Tris (pH 7.4). The remaining marrow cells were washed 3x with MEM and transfected for 2 hours with 9 µg of pCMV-LMP-1s (forward or reverse orientation) per 3 x 10<sup>6</sup> cells. The transfected cells were then washed 2X with MEM and resuspended at a concentration of 3 x 10<sup>7</sup> cells/ml.

25 The cell suspension (100 µl) was applied via sterile pipette to a sterile 2 x 5 mm type I bovine collagen disc (Sulzer Orthopaedics, Wheat Ridge, CO). The discs were surgically implanted subcutaneously on the skull, chest, abdomen or dorsal spine of 4-5 week old athymic rats (rnu/rnu). The animals were scarified at 3-4 weeks, at which time the discs or surgical areas were  
30 excised and fixed in 70% ethanol. The fixed specimens were analyzed by radiography and undecalcified histologic examination was performed on 5 m

thick sections stained with Goldner Trichrome. Experiments were also performed using devitalized (guanidine extracted) demineralized bone matrix (Osteotech, Shrewsbury, NJ) in place of collagen discs.

Radiography revealed a high level of mineralized bone formation that conformed to the form of the original collagen disc containing LMP-1s transfected marrow cells. No mineralized bone formation was observed in the negative control (cells transfected with a reverse-oriented version of the LMP-1s cDNA that did not code for a translated protein), and absorption of the carrier appeared to be well underway.

Histology revealed new bone trabeculae lined with osteoblasts in the LMP-1s transfected implants. No bone was seen along with partial resorption of the carrier in the negative controls.

Radiography of a further experiment in which 18-sets (9 negative control pCMV-LMP-REV & 9 experimental pCMV-LMP-1s) of implants were added to sites alternating between lumbar and thoracic spine in athymic rats demonstrated 0/9 negative control implants exhibiting bone formation (spine fusion) between vertebrae. All nine of the pCMV-LMP-1s treated implants exhibited solid bone fusions between vertebrae.

EXAMPLE 26: The Synthesis of pHIS-5' ATG LMP-1s Expression Vector from the sequences Demonstrated in Examples 2 and 3.

The 717 base-pair clone (Example 17) was digested with Clal and EcoRV (New England Biologicals, city, MA). A small fragment (~250 base-pairs) was gel purified. Clone No. 7 (Example 18) was digested with Clal and XbaI. A 1400 base-pair fragment was gel purified from that digest. The isolated 250 base-pair and 1400 base-pair cDNA fragments were ligated by standard methods to form a fragment of ~1650 bp. The pHIS-A vector (InVitrogen) was digested with EcoRV and XbaI. The linearized vector was recovered and ligated to the chimeric 1650 base-pair cDNA fragment. The ligated product was cloned and amplified by standard methods, and the pHIS-

A-5' ATG LMP-1s expression vector, also denominated as the vector pHis-A with insert HLMP-1s, was deposited at the ATCC as previously described.

EXAMPLE 27: The Induction of Bone Nodule Formation and Mineralization *In*

5 *Vitro* With pHis-5' ATG LMP-1s Expression Vector

Rat calvarial cells were isolated and grown in secondary culture according to Example 1. Cultures were either unstimulated or stimulated with glucocorticoid (GC) according to Example 1. The cultures were transfected  
10 with 3 g of recombinant pHis-A vector DNA/well as described in Example 25. Mineralized nodules were visualized by Von Kossa staining according to Example 3.

Human LMP-1s gene product overexpression alone (*i.e.*, without GC stimulation) induced significant bone nodule formation (~203 nodules/well) *in*  
15 *vitro*. This is approximately 50% of the amount of nodules produced by cells exposed to the GC positive control (~412 nodules/well). Similar results were obtained with cultures transfected with pHisA-LMP-Rat Expression vector (~152 nodules/well) and pCMV2/LMP-Rat-Fwd (~206 nodules/well). In contrast, the negative control pCMV2/LMP-Rat-Rev yielded (~2 nodules/well),  
20 while approximately 4 nodules/well were seen in the untreated plates. These data demonstrate that the human LMP-1 cDNA was at least as osteoinductive as the rat LMP-1 cDNA in this model system. The effect in this experiment was less than that observed with GC stimulation; but in some the effect was comparable.

25

EXAMPLE 28: LMP Induces Secretion of a Soluble Osteoinductive Factor

Overexpression of RLMP-1 or HLMP-1s in rat calvarial osteoblast cultures as described in Example 24 resulted in significantly greater nodule  
30 formation than was observed in the negative control. To study the mechanism of action of LIM mineralization protein conditioned medium was harvested at

different time points, concentrated to 10 X, sterile filtered, diluted to its original concentration in medium containing fresh serum, and applied for four days to untransfected cells.

Conditioned media harvested from cells transfected with RLMP-1 or  
5 HLMP-1s at day 4 was approximately as effective in inducing nodule formation as direct overexpression of RLMP-1 in transfected cells. Conditioned media from cells transfected with RLMP-1 or HLMP-1 in the reverse orientation had no apparent effect on nodule formation. Nor did conditioned media harvested from LMP-1 transfected cultures before day 4 induce nodule formation. These  
10 data suggest that expression of LMP-1 caused the synthesis and/or secretion of a soluble factor, which did not appear in culture medium in effective amounts until 4 days post transfection.

Since overexpression of rLMP-1 resulted in the secretion of an osteoinductive factor into the medium, Western blot analysis was used to  
15 determine if LMP-1 protein was present in the medium. The presence of rLMP-1 protein was assessed using antibody specific for LMP-1 (QDPDEE) and detected by conventional means. LMP-1 protein was found only in the cell layer of the culture and not detected in the medium.

Partial purification of the osteoinductive soluble factor was accomplished  
20 by standard 25% and 100% ammonium sulfate cuts followed by DE-52 anion exchange batch chromatography (100 mM or 500 mM NaCl). All activity was observed in the high ammonium sulfate, high NaCl fractions. Such localization is consistent with the possibility of a single factor being responsible for conditioning the medium.

25

#### EXAMPLE 29: Gene Therapy In Lumbar Spine Fusion Mediated by Low Dose Adenovirus

This study determined the optimal dose of adenoviral delivery of the  
30 LMP-1 cDNA (SEQ ID NO: 2) to promote spine fusion in normal, that is, immune competent, rabbits.



A replication-deficient human recombinant adenovirus was constructed with the LMP-1 cDNA (SEQ ID NO: 2) driven by a CMV promoter using the Adeno-Quest™ Kit (Quantum Biotechnologies, Inc., Montreal). A commercially available (Quantum Biotechnologies, Inc., Montreal) recombinant adenovirus containing the beta-galactosidase gene was used as a control.

Initially, an *in vitro* dose response experiment was performed to determine the optimal concentration of adenovirus-delivered LMP-1 ("AdV-LMP-1") to induce bone differentiation in rat calvarial osteoblast cultures using a 60-minute transduction with a multiplicity of infection ("MOI") of 0.025, 0.25, 2.5, or 25 plaque-forming units (pfu) of virus per cell. Positive control cultures were differentiated by a 7-day exposure to  $10^{-9}$  M glucocorticoid ("GC"). Negative control cultures were left untreated. On day 14, the number of mineralized bone nodules was counted after von Kossa staining of the cultures, and the level of osteocalcin secreted into the medium (pmol/mL) was measured by radioimmunoassay (mean  $\pm$  SEM).

The results of this experiment are shown in Table I. Essentially no spontaneous nodules formed in the untreated negative control cultures. The data show that a MOI equal to 0.25 pfu/cell is most effective for osteoinducing bone nodules, achieving a level comparable to the positive control (GC). Lower and higher doses of adenovirus were less effective.

TABLE I

Outcome	AdV-LMP-1 Dose (MOI)					
	Neg. Ctrl.	GC	0.025	0.25	2.5	25
Bone Nodules	0.5 $\pm$ 0.2	188 $\pm$ 35	79.8 $\pm$ 13	145.1 $\pm$ 13	26.4 $\pm$ 15	87.6 $\pm$ 2
Osteocalcin	1.0 $\pm$ 0.1	57.8 $\pm$ 9	28.6 $\pm$ 11	22.8 $\pm$ 1	18.3 $\pm$ 3	26.0 $\pm$ 2

*In vivo* experiments were then performed to determine if the optimal *in vitro* dose was capable of promoting intertransverse process spine fusions in skeletally mature New Zealand white rabbits. Nine rabbits were anesthetized and 3 cc of bone marrow was aspirated from the distal femur through the intercondylar notch using an 18 gauge needle. The buffy coat was then isolated, a 10-minute transduction with AdV-LMP-1 was performed, and the cells were returned to the operating room for implantation. Single level posterolateral lumbar spine arthrodesis was performed with decortication of transverse processes and insertion of carrier (either rabbit devitalized bone matrix or a collagen sponge) containing 8-15 million autologous nucleated buffy coat cells transduced with either AdV-LMP-1 (MOI = 0.4) or AdV-BGal (MOI = 0.4). Rabbits were euthanized after 5 weeks and spine fusions were assessed by manual palpation, plain x-rays, CT scans, and undecalcified histology.

The spine fusion sites that received AdV-LMP-1 induced solid, continuous spine fusion masses in all nine rabbits. In contrast, the sites receiving AdV-BGal, or a lower dose of AdV-LMP-1 (MOI = 0.04) made little or no bone and resulted in spine fusion at a rate comparable to the carrier alone (< 40%). These results were consistent as evaluated by manual palpation, CT scan, and histology. Plain radiographs, however, sometimes overestimated the amount of bone that was present, especially in the control sites. LMP-1 cDNA delivery and bone induction was successful with both of the carrier materials tested. There was no evidence of systemic or local immune response to the adenovirus vector.

These data demonstrate consistent bone induction in a previously validated rabbit spine fusion model which is quite challenging. Furthermore, the protocol of using autogenous bone marrow cells with intraoperative *ex vivo* gene transduction (10 minutes) is a more clinically feasible procedure than other methods that call for overnight transduction or cell expansion for weeks in culture. In addition, the most effective dose of recombinant adenovirus (MOI=0.25) was substantially lower than doses reported in other gene therapy

applications (MOI-40-500). We believe this is due to the fact that LMP-1 is an intracellular signaling molecule and may have powerful signal amplification cascades. Moreover, the observation that the same concentration of AdV-LMP-1 that induced bone in cell culture was effective *in vivo* was also surprising given the usual required increase in dose of other growth factors when translating from cell culture to animal experiments. Taken together, these observations indicate that local gene therapy using adenovirus to deliver the LMP-1 cDNA is possible and the low dose required will likely minimize the negative effects of immune response to the adenovirus vector.

EXAMPLE 30: Use of Peripheral Venous Blood Nucleated Cells (Buffy Coat) for Gene Therapy With LMP-1 cDNA To Make Bone

In four rabbits we performed spine fusion surgery as above (Example 29) except the transduced cells were the buffy coat from venous blood rather than bone marrow. These cells were transfected with Adeno-LMP or pHIS-LMP plasmid and had equivalent successful results as when bone marrow cells were used. This discovery of using ordinary venous blood cells for gene delivery makes gene therapy more feasible clinically since it avoids painful marrow harvest under general anesthesia and yields two times more cells per mL of starting material.

EXAMPLE 31: Isolation of Human LMP-1 Splice Variants

Intron/Exon mRNA transcript splice variants are a relatively common regulatory mechanism in signal transduction and cellular/tissue development. Splice variants of various genes have been shown to alter protein-protein, protein-DNA, protein-RNA, and protein-substrate interactions. Splice variants may also control tissue specificity for gene expression allowing different forms (and therefore functions) to be expressed in various tissues. Splice variants are a common regulatory phenomenon in cells. It is possible that the LMP

splice variants may result in effects in other tissues such as nerve regeneration, muscle regeneration, or development of other tissues.

To screen a human heart cDNA library for splice variants of the HLMP-1 sequence, a pair of PCR primer corresponding to sections of SEQ ID NO: 22 was prepared. The forward PCR primer, which was synthesized using standard techniques, corresponds to nucleotides 35-54 of SEQ ID NO: 22. It has the following sequence:

5' GAGCCGGCATCATGGATTCC 3' (SEQ ID NO: 35)

10

The reverse PCR primer, which is the reverse complement of nucleotides 820-839 in SEQ ID NO: 22, has the following sequence:

5' GCTGCCTGCACAATGGAGGT 3' (SEQ ID NO: 36)

15

The forward and reverse PCR primers were used to screen human heart cDNA (ClonTech, Cat No. 7404-1) for sequences similar to HLMP-1 by standard techniques, using a cycling protocol of 94°C for 30 seconds, 64°C for 30 seconds, and 72°C for 1 minute, repeated 30 times and followed by a 10 minute incubation at 72°C. The amplification cDNA sequences were gel-purified and submitted to the Emory DNA Sequence Core Facility for sequencing. The clones were sequenced using standard techniques and the sequences were examined with PCGENE (Intelligenetics; Programs SEQUIN and NALIGN) to determine homology to SEQ ID NO: 22. Two homologous nucleotide sequences with putative alternative splice sites compared to SEQ ID NO: 22 were then translated to their respective protein products with Intelligenetic's program TRANSL.

25

One of these two novel human cDNA sequences (SEQ ID NO: 37) comprises 1456 bp:

30

10 20 30 40 50 60

CGACGCAGAG CAGCGCCCTG GCCGGGCCAA GCAGGAGCCG GCATCATGGA TTCCTTCAAG  
 70 80 90 100 110 120  
 GTAGTGCTGG AGGGGCCAGC ACCTTGGGGC TTCCGGCTGC AAGGGGGCAA GGAATTCAAT  
 130 140 150 160 170 180  
 5 GTGCCCCCTCT CCATTTCCCG GCTCACTCCT GGGGGCAAAG CGGCGCAGGC CGGAGTGGCC  
 190 200 210 220 230 240  
 GTGGGTGACT GGGTGCTGAG CATCGATGGC GAGAATGCGG GTAGCCTCAC ACACATCGAA  
 250 260 270 280 290 300  
 GCTCAGAACA AGATCCGGGC CTGCGGGGAG CGCCTCAGCC TGGGCCTCAG CAGGGCCCAG  
 10 310 320 \* 330 340 \* 350 360  
 CCGGTTTCAGA GCAAACCGCA GAAGGTGCAG ACCCCTGACA AACAGCCGCT CCGACCGCTG  
 370 380 390 400 410 420  
 GTCCCAGATG CCAGCAAGCA GCGGCTGATG GAGAACACAG AGGACTGGCG GCCGCGGCCG  
 430 440 450 460 470 480  
 15 GGGACAGGCC AGTCGCGTTC CTTCCGCATC CTTGCCCACC TCACAGGCAC CGAGTTCATG  
 490 500 510 520 530 540  
 CAAGACCCGG ATGAGGAGCA CCTGAAGAAA TCAAGCCAGG TGCCCAGGAC AGAAGCCCCA  
 550 560 570 580 590 600  
 GCCCCAGCCT CATCTACACC CCAGGAGCCC TGGCCTGGCC CTACCGCCCC CAGCCCTACC  
 20 610 620 630 640 650 660  
 AGCCGCCCCG CCTGGGCTGT GGACCCTGCG TTTGCCGAGC GCTATGCCCC GGACAAAACG  
 670 680 690 700 710 720  
 AGCACAGTGC TGACCCGGCA CAGCCAGCCG GCCACGCCCA CGCCGCTGCA GAGCCGCACC  
 730 740 750 760 770 780  
 25 TCCATTGTGC AGGCAGCTGC CGGAGGGGTG CCAGGAGGGG GCAGCAACAA CGGCAAGACT  
 790 800 810 820 830 840  
 CCCGTGTGTC ACCAGTGCCA CAAGGTCATC CGGGGCCGCT ACCTGGTGGC GTTGGGCCAC  
 850 860 870 880 890 900

```

          GCGTACCACC CGGAGGAGTT TGTGTGTAGC CAGTGTGGGA AGGTCCTGGA AGAGGGTGGC
          910          920          930          940          950          960
          TTCTTTGAGG AGAAGGGCGC CATCTTCTGC CCACCATGCT ATGACGTGCG CTATGCACCC
          970          980          990          1000          1010          1020
5          AGCTGTGCCA AGTGCAAGAA GAAGATTACA GGCGAGATCA TGCACGCCCT GAAGATGACC
          1030          1040          1050          1060          1070          1080
          TGGCACGTGC ACTGCTTTAC CTGTGCTGCC TGCAAGACGC CCATCCGGAA CAGGGCCTTC
          1090          1100          1110          1120          1130          1140
          TACATGGAGG AGGGCGTGCC CTATTGCGAG CGAGACTATG AGAAGATGTT TGGCACGAAA
10          1150          1160          1170          1180          1190          1200
          TGCCATGGCT GTGACTTCAA GATCGACGCT GGGGACCGCT TCCTGGAGGC CCTGGGCTTC
          1210          1220          1230          1240          1250          1260
          AGCTGGCATG ACACCTGCTT CGTCTGTGCG ATATGTCAGA TCAACCTGGA AGGAAAGACC
          1270          1280          1290          1300          1310          1320
15          TTCTACTCCA AGAAGGACAG GCCTCTCTGC AAGAGCCATG CCTTCTCTCA TGTGTGAGCC
          1330          1340          1350          1360          1370          1380
          CCTTCTGCCC ACAGCTGCCG CGGTGGCCCC TAGCCTGAGG GGCCTGGAGT CGTGGCCCTG
          1390          1400          1410          1420          1430          1440
          CATTTCTGGG TAGGGCTGGC AATGGTTGCC TTAACCCTGG CTCCTGGCCC GAGCCTGGGC
20          1450
          TCCCGGGCCC TGCCCA

```

Reading frame shifts caused by the deletion of a 119 bp fragment  
 (between \*) and the addition of a 17 bp fragment (underlined) results in a  
 25 truncated gene product having the following derived amino acid sequence  
 (SEQ ID NO: 38):

```

          10          20          30          40          50          60

```

```

MDSFKVVLEG PAPWGFRLQG GKDFNVPLSI SRLTPGGKAA QAGVAVGDWV LSIDGENAGS
      70          80          90          100          110          120
LTHIEAQNKI RACGERLSLG LSRAQPVQNK PQKVQTPDKQ PLRPLVPDAS KQRLMENTED
      130          140          150          160          170          180
5  WRPRPGTGQS RSFRILAHLT GTEFMQDPDE EHLKKSSQVP RTEAPAPASS TPQEPWPGPT
      190          200          210          220          230          240
APSPTSPPW AVDPFAERY APDKTSTVLT RHSQPATPTP LQSRTSIVQA AAGGVPGGGS
      250          260          270          280          290          300
NNGKTPVCHQ CHQVIRARYL VALGHAYHPE EFVCSQCGKV LEEGGFFEEK GAIFCPPCYD
10      310          320          330          340          350          360
VRYAPSCAKC KKKITGEIMH ALKMTWHVLC FTCAACKTPI RNRAFYMEEG VPYCERDYEK
      370          380          390          400          410          420
MFGTKCQWCD FKIDAGDRFL EALGFSWHD TCFVCAICQIN LEGKTFYSKK DRPLCKSHAF
SHV

```

- 15           This 423 amino acid protein demonstrates 100% homology to the protein shown in Sequence ID No. 10, except for the sequence in the highlighted area (amino acids 94-99), which are due to the nucleotide changes depicted above.

          The second novel human heart cDNA sequence (SEQ ID NO: 39)  
20 comprises 1575 bp:

```

      10          20          30          40          50          60
CGACGCAGAG CAGCGCCCTG GCCGGGCCAA GCAGGAGCCG GCATCATGGA TTCCTTCAAG
      70          80          90          100          110          120
25 GTAGTGCTGG AGGGGCCAGC ACCTTGGGGC TTCCGGCTGC AAGGGGGCAA GGA CTTCAAT
      130          140          150          160          170          180
GTGCCCCTCT CCATTTCCCG GCTCACTCCT GGGGGCAAAG CGGCGCAGGC CGGAGTGGCC
      190          200          210          220          230          240

```

GTGGGTGACT GGGTGCTGAG CATCGATGGC GAGAATGCGG GTAGCCTCAC ACACATCGAA  
 250 260 270 280 290 300  
 GCTCAGAACA AGATCCGGGC CTGCGGGGAG CGCCTCAGCC TGGGCCTCAG CAGGGCCCAG  
 310 320 330 340 350 360  
 5 CCGGTTTCTCAGA GCAAACCGCA GAAGGCCTCC GCGGACCTCC GCGGTACACC  
 370 380 390 400 410 420  
 TTTGCACCCA GCGTCTCCCT CAACAAGACG GCGGCGCCCT TTGGGGCGCC CCCGCCCCGCT  
 430 440 450 460 470 480  
 GACAGCGCCC CGCAACAGAA TGGGTGAGAG CCCCTGACAA ACAGCCGCTC CGACCGCTGG  
 10 490 500 510 520 530 540  
 TCCCAGATGC CAGCAAGCAG CGGCTGATGG AGAACACAGA GGACTGGCGG CCGCGGCCGG  
 550 560 570 580 590 600  
 GGACAGGCCA GTCGCGTTCC TTCCGCATCC TTGCCCACCT CACAGGCACC GAGTTCATGC  
 610 620 630 640 650 660  
 15 AAGACCCGGA TGAGGAGCAC CTGAAGAAAT CAAGCCAGGT GCCCAGGACA GAAGCCCCAG  
 670 680 690 700 710 720  
 CCCCAGCCTC ATCTACACCC CAGGAGCCCT GGCCTGGCCC TACCGCCCCC AGCCCTACCA  
 730 740 750 760 770 780  
 GCCGCCCCGCC CTGGGCTGTG GACCCTGCGT TTGCCGAGCG CTATGCCCCG GACAAAACGA  
 20 790 800 810 820 830 840  
 GCACAGTGCT GACCCGGCAC AGCCAGCCGG CCACGCCCAC GCCGCTGCAG AGCCGCACCT  
 850 860 870 880 890 900  
 CCATTGTGCA GGCAGCTGCC GGAGGGGTGC CAGGAGGGGG CAGCAACAAC GGCAAGACTC  
 910 920 930 940 950 960  
 25 CCGTGTGTCA CCAGTGCCAC AAGGTCATCC GGGGCCGCTA CCTGGTGGCG TTGGGCCACG  
 970 980 990 1000 1010 1020  
 CGTACCACCC GGAGGAGTTT GTGTGTAGCC AGTGTGGGAA GGTCTTGGA GAGGGTGGCT  
 1030 1040 1050 1060 1070 1080



TCTTTGAGGA GAAGGGCGCC ATCTTCTGCC CACCATGCTA TGACGTGCGC TATGCACCCA  
 1090 1100 1110 1120 1130 1140  
 GCTGTGCCAA GTGCAAGAAG AAGATTACAG GCGAGATCAT GCACGCCCTG AAGATGACCT  
 1150 1160 1170 1180 1190 1200  
 5 GGCACGTGCA CTGCTTTACC TGTGCTGCCT GCAAGACGCC CATCCGGAAC AGGGCCTTCT  
 1210 1220 1230 1240 1250 1260  
 ACATGGAGGA GGGCGTGCCC TATTGCGAGC GAGACTATGA GAAGATGTTT GGCACGAAAT  
 1270 1280 1290 1300 1310 1320  
 GCCATGGCTG TGA~~CTT~~CAAG ATCGACGCTG GGGACCGCTT CCTGGAGGCC CTGGGCTTCA  
 10 1330 1340 1350 1360 1370 1380  
 GCTGGCATGA CACCTGCTTC GTCTGTGCGA TATGTCAGAT CAACCTGGAA GGAAAGACCT  
 1390 1400 1410 1420 1430 1440  
 TCTACTCAA GAAGGACAGG CCTCTCTGCA AGAGCCATGC CTTCTCTCAT GTGTGAGCCC  
 1450 1460 1470 1480 1490 1500  
 15 CTTCTGCCCCA CAGCTGCCGC GGTGGCCCCCT AGCCTGAGGG GCCTGGAGTC GTGGCCCTGC  
 1510 1520 1530 1540 1550 1560  
 ATTTCTGGGT AGGGCTGGCA ATGGTTGCCT TAACCCTGGC TCCTGGCCCG AGCCTGGGCT  
 1570  
 CCCGGGCCCT GCCCA

20 Reading frame shifts caused by the addition of a 17 bp fragment  
 (bolded, italicized and underlined) results in an early translation stop codon at  
 position 565-567 (underlined).

The derived amino acid sequence (SEQ ID NO: 40) consists of 153  
 amino acids:

25

10 20 30 40 50 60  
 MDSFKVVLEG PAPWGFRLQG GKDFNVPLSI SRLTPGGKAA QAGVAVGDWV LSIDGENAGS  
 70 80 90 100 110 120

LTHIEAQNKI RACGERLSLG LSRAQPVQSK POKASAPAD PPRYTFAPSV SLNKTARFPG

130

140

150

APPPADSAPO ONGCRPLTNS RSDRWSOMPA SSG

- 5 This protein demonstrates 100% homology to SEQ ID NO: 10 until amino acid 94, where the addition of the 17 bp fragment depicted in the nucleotide sequence results in a frame shift. Over amino acids 94-153, the protein is not homologous to SEQ ID NO: 10. Amino acids 154-457 in SEQ ID NO: 10 are not present due to the early stop codon depicted in nucleotide sequence.

EXAMPLE 32: Genomic HLMP-1 Nucleotide Sequence

Applicants have identified the genomic DNA sequence encoding HLMP-1, including putative regulatory elements associated with HLMP-1 expression.

- 5 The entire genomic sequence is shown in SEQ ID. NO: 41. This sequence was derived from AC023788 (clone RP11-564G9), Genome Sequencing Center, Washington University School of Medicine, St. Louis, MO.

- The putative promoter region for HLMP-1 spans nucleotides 2,660-8,733 in SEQ ID NO: 41. This region comprises, among other things, at least  
 10 ten potential glucocorticoid response elements ("GREs") (nucleotides 6148-6153, 6226-6231, 6247-6252, 6336-6341, 6510-6515, 6552-6557, 6727-6732, 6752-6757, 7738-7743, and 8255-8260), twelve potential Sma-2 homologues to Mothers against Drosophilla decapentaplegic ("SMAD") binding element sites (nucleotides 3569-3575, 4552-4558, 4582-4588, 5226-5232, 6228-6234,  
 15 6649-6655, 6725-6731, 6930-6936, 7379-7384, 7738-7742, 8073-8079, and 8378-8384), and three TATA boxes (nucleotides 5910-5913, 6932-6935, and 7380-7383). The three TATA boxes, all of the GREs, and eight of the SMAD binding elements ("SBEs") are grouped in the region spanning nucleotides 5,841-8,733 in SEQ ID NO: 41. These regulatory elements can be used, for  
 20 example, to regulate expression of exogenous nucleotide sequences encoding proteins involved in the process of bone formation. This would permit systemic administration of therapeutic factors or genes relating to bone formation and repair, as well as factors or genes associated with tissue differentiation and development.

- 25 In addition to the putative regulatory elements, 13 exons corresponding to the nucleotide sequence encoding HLMP-1 have been identified. These exons span the following nucleotides in SEQ ID NO: 41:

30	Exon 1	8733-8767
	Exon 2	9790-9895
	Exon 3	13635-13787

	Exon 4	13877-13907
	Exon 5	14387-14502
	Exon 6	15161-15297
	Exon 7	15401-15437
5	Exon 8	16483-16545
	Exon 9	16689-16923
	Exon 10	18068-18248
	Exon 11	22117-22240
	Exon 12	22323-22440
10	Exon 13	22575-22911

In HLMP-2 there is another exon (Exon 5A), which spans nucleotides 14887-14904.

All cited publications and patents are hereby incorporated by reference in their entirety.


15        While the foregoing specification teaches the principles of the present invention, with examples provided for the purpose of illustration, it will be appreciated by one skilled in the art from reading this disclosure that various changes in form and detail can be made without departing from the true scope of the invention.

Applicant's or agent's file reference number 6148.. 2-304	International application PCT/US00/11664
--	---

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>17</u> , line <u>20</u>	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Blvd. Manassas, VA 20110-2209	
Date of deposit 14 April 2000	Accession Number PTA-1698
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<b>For receiving Office use only</b>	<b>For International Bureau use only</b>
<input type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer	17 JUL 2000
	Authorized officer 

Applicant's or agent's file reference number 6148...2-304	International application PCT/US00/11664
--	--

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 17, line 21	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Blvd. Manassas, VA 20110-2209	
Date of deposit 14 April 2000	Accession Number PTA-1699
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only
<input type="checkbox"/> This sheet was received with the international application
Authorized officer

For International Bureau use only
<input type="checkbox"/> This sheet was received by the International Bureau on: 17 JUL 2000
Authorized officer

We claim:

1. An isolated nucleic acid molecule comprising a SEQ ID NO: 37 or  
SEQ ID NO: 39.
2. An isolated human LMP protein encoded by SEQ ID NO: 37 or SEQ  
5 ID NO: 39.
3. A vector comprising the isolated nucleic acid molecule of claims 1.
4. A host cell comprising the vector of claim 3.
5. The host cell of claim 4, wherein the host cell is selected from the  
group consisting of prokaryotic cells, yeast cells and mammalian  
10 cells
6. The isolated nucleic acid molecule of claim 1, further comprising a  
label.
7. A human LIM mineralization protein comprising an amino acid  
sequence selected from the group consisting of SEQ ID NO: 38 and SEQ ID  
15 NO: 40.
8. A monoclonal antibody specific for a HLMP-2 (SEQ ID NO: 38) or  
HLMP-3 (SEQ ID NO: 40).
9. A method of inducing bone formation comprising transfecting  
osteogenic precursor cells or peripheral blood leukocytes with an isolated  
20 nucleic acid molecule comprising SEQ ID NO: 39.
10. The method of claim 9, wherein the isolated nucleic acid molecule is  
in a vector.
11. The method of claim 10, wherein the vector is an expression vector.
12. The method of claim 11, wherein the vector is a plasmid.
- 25 13. The method of claim 11, wherein the vector is a virus.
14. The method of claim 13, wherein the virus is an adenovirus
15. The method of claim 13, wherein the virus is a retrovirus.
16. The method of claim 9, wherein the osteogenic precursor cells or  
peripheral blood leukocytes are transfected *ex vivo*.
- 30 17. The method of claim 9, wherein the osteogenic precursor cells are  
transfected *in vivo* by direct injection of the isolated nucleic acid molecule.

18. The method of claim 9, wherein the LIM mineralization protein is HLMP-3 (SEQ ID NO: 40).

19. A method of fusing a spine, comprising:

- (a) transfecting osteogenic precursor cells or peripheral blood leukocytes with an isolated nucleic acid molecule comprising SEQ ID NO: 39;
- (b) admixing the transfected osteogenic precursor cells or peripheral blood leukocytes with a matrix; and
- (c) contacting the matrix with the spine;

wherein expression of the nucleotide sequence causes mineralized bone formation in the matrix.

20. The method of claim 19, wherein the osteogenic precursor cells or peripheral blood cells are transfected *ex vivo*.

21. A method of inducing systemic bone formation in a mammalian host in need thereof, comprising:

- a) transfecting osteogenic precursor cells or peripheral blood leukocytes with a vector that is stably maintained in the osteogenic precursor cells or peripheral blood leukocytes, the vector comprising SEQ ID NO: 39 operably linked to a regulatable promoter, wherein the regulatable promoter responds to an exogenous control compound; and
- (b) administering to the host, as needed, an amount of the exogenous control substance effective to cause expression of SEQ ID NO: 39.



22. A method of stimulating production of an osteogenic soluble factor by an osteogenic cell, comprising:

(a) transfecting the osteogenic cell or peripheral blood leukocyte with an isolated nucleic acid molecule comprising SEQ ID NO: 39; and

5 (b) overexpressing the isolated nucleic acid molecule.

23. An osteogenic soluble factor produced by the method of claim 22.

24. The osteogenic soluble factor of claim 23, wherein the osteogenic factor is a protein.

25. A method of inhibiting the expression of HLMP-2 or HLMP-3  
10 comprising transfecting a cell wherein HLMP-2 or HLMP-3 is expressed with an antisense oligonucleotide.

26. The method of claim of claim 17, wherein the isolated nucleic acid molecule is in a vector selected from the group consisting of a plasmid and a virus.

15 27. The method of claim 26, wherein the vector is a plasmid, which plasmid is directly injected into muscle tissue.

28. A method of inhibiting bone formation comprising transfecting osteogenic precursor cells or peripheral blood leukocytes with an isolated nucleic acid molecule comprising SEQ ID NO: 37.

20 29. The method of claim 28, wherein the isolated nucleic acid molecule is in a vector.

30. The method of claim 29, wherein the vector is an expression vector.

31. The method of claim 30, wherein the vector is a plasmid.

32. The method of claim 30, wherein the vector is a virus.

25 33. The method of claim 32, wherein the virus is an adenovirus

34. The method of claim 32, wherein the virus is a retrovirus.

35. The method of claim 28, wherein the osteogenic precursor cells or peripheral blood leukocytes are transfected *ex vivo*.

30 36. The method of claim 28, wherein the osteogenic precursor cells are transfected *in vivo* by direct injection of the isolated nucleic acid molecule.

37. The method of claim 28, wherein the LIM mineralization protein is HLMP-2 (SEQ ID NO: 38).

## SEQUENCE LISTING

<110> Boden M.D., Scott D  
Hair Ph.D., Gregory A

<120> LIM Mineralization Protein Splice Variants

<130> splice variant LMP

<140> 6148.0222-00304

<141> 2000-04-28

<150> 60/132,021

<151> 1999-04-30

<160> 42

<170> PatentIn Ver. 2.1

<210> 1

<211> 457

<212> PRT

<213> Rattus norvegicus

<400> 1

Met	Asp	Ser	Phe	Lys	Val	Val	Leu	Glu	Gly	Pro	Ala	Pro	Trp	Gly	Phe
1				5					10					15	

Arg	Leu	Gln	Gly	Gly	Lys	Asp	Phe	Asn	Val	Pro	Leu	Ser	Ile	Ser	Arg
		20						25					30		

Leu	Thr	Pro	Gly	Gly	Lys	Ala	Ala	Gln	Ala	Gly	Val	Ala	Val	Gly	Asp
		35					40						45		

Trp	Val	Leu	Ser	Ile	Asp	Gly	Glu	Asn	Ala	Gly	Ser	Leu	Thr	His	Ile
	50					55					60				

Glu	Ala	Gln	Asn	Lys	Ile	Arg	Ala	Cys	Gly	Glu	Arg	Leu	Ser	Leu	Gly
65					70					75				80	

Leu	Ser	Arg	Ala	Gln	Pro	Ala	Gln	Ser	Lys	Pro	Gln	Lys	Ala	Leu	Thr
			85						90					95	

Pro	Pro	Ala	Asp	Pro	Pro	Arg	Tyr	Thr	Phe	Ala	Pro	Ser	Ala	Ser	Leu
			100					105					110		

Asn	Lys	Thr	Ala	Arg	Pro	Phe	Gly	Ala	Pro	Pro	Pro	Thr	Asp	Ser	Ala
		115					120						125		

Leu	Ser	Gln	Asn	Gly	Gln	Leu	Leu	Arg	Gln	Leu	Val	Pro	Asp	Ala	Ser
	130					135					140				

Lys	Gln	Arg	Leu	Met	Glu	Asn	Thr	Glu	Asp	Trp	Arg	Pro	Arg	Pro	Gly
145					150					155					160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175  
 Glu Phe Met Gln Asp Pro Asp Glu Glu Phe Met Lys Lys Ser Ser Gln  
 180 185 190  
 Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Thr Ile Pro Gln Glu  
 195 200 205  
 Ser Trp Pro Gly Pro Thr Thr Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220  
 Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240  
 Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255  
 Asn Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Gly Thr Gly Gly  
 260 265 270  
 Gly Ser Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Lys Ile  
 275 280 285  
 Ile Arg Gly Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu  
 290 295 300  
 Glu Phe Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe  
 305 310 315 320  
 Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro Ser Cys Tyr Asp Val Arg  
 325 330 335  
 Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile  
 340 345 350  
 Met His Ala Leu Lys Met Thr Trp His Val Pro Cys Phe Thr Cys Ala  
 355 360 365  
 Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly  
 370 375 380  
 Ala Pro Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys  
 385 390 395 400  
 Arg Gly Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala  
 405 410 415  
 Leu Gly Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln  
 420 425 430  
 Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Lys Pro Leu  
 435 440 445  
 Cys Lys Ser His Ala Phe Ser His Val  
 450 455

<210> 2  
 <211> 1696  
 <212> DNA  
 <213> *Rattus norvegicus*

<400> 2  
 gcacgaggat cccagcgcggt ctccctggagg ccgccaggca gccgcccagc cggggcattca 60  
 ggagcaggta ccatggattc cttcaaggta gtgctggagg gacctgcccc ttggggcttc 120  
 cgtctgcaag ggggcaagga cttcaacgtg cccctctcca tctctcggct cactcctgga 180  
 ggcaaggccg cacaggccgg tgtggccgtg ggagactggg tactgagtat cgacgggtgag 240  
 aacgcccga gacctacaca cattgaagcc cagaacaaga tccgtgcttg tggggagcgc 300  
 ctgagcctgg gtcttagcag agcccagcct gctcagagca aaccacagaa ggccctgacc 360  
 cctcccgccg acccccagag gtacactttt gcaccaagcg cctccctcaa caagacggcc 420  
 cggcccttcg gggcaccccc acctactgac agcgccctgt cgcagaatgg acagctgctc 480  
 agacagctgg tccctgatgc cagcaagcag cggctgatgg agaatactga agactggcgc 540  
 ccgcgccag ggacaggcca gtcccgttcc ttccgcatcc ttgctcacct caccgggcaca 600  
 gagttcatgc aagaccggga tgaggaattc atgaagaagt caagccaggt gccaggaca 660  
 gaagccccag cccagcctc aaccataccc caggaatcct ggctggccc caccacccc 720  
 agccccacca gccgcccacc ctgggeegta gatcctgcat ttgctgagcg ctatgcccc 780  
 gacaaaacca gcacagtgtc gaccgcagac agccagccag ccacacctac gcctctgcag 840  
 aaccgcacct ccatagttca ggctgcagct ggagggggca caggaggagg cagcaacaat 900  
 ggcaagacgc ctgtatgcca ccagtgccac aagatcatcc gccgcccata cctggtagca 960  
 ctggggccacg cgtaccatcc tgaggaatct gtgtgcagcc agtgtgggaa ggtcctggaa 1020  
 gaggggtggt tcttcgagga gaaggagct atcttttgcc cctcctgcta tgatgtgcgc 1080  
 tatgcaccca gctgtgcca atgcaagaag aagatcactg gagagatcat gcatgcgctg 1140  
 aagatgacct ggcatgttcc ctgcttcacc tgtgcagcct gcaaaacccc tatccgcaac 1200  
 agggctttct acatggagga gggggctccc tactgcgagc gagattacga gaagatgttt 1260  
 ggcacaaagt gtcgcggtg tgacttcaag atcgatgccg gggaccgttt cctggaagcc 1320  
 ctgggtttca gctggcatga tacgtgtttt gtttgcgcaa tatgtcaaat caacttgga 1380  
 ggaaagacct tctactccaa gaaggacaag cccctgtgca agagccatgc cttttccac 1440  
 gtatgagcac ctccacacac tactgccacc ctactctgcc agaagggtga taaaatgaga 1500  
 gagctctctc tcctcgacc tttctgggtg gggtggcag ccattgtcct agccttggt 1560  
 cctggccaga tctggggct cctcctcac agtccccctt cccacacttc ctccaccacc 1620  
 accaccgtca ctcacaggtg ctgacctcct agccccagtt cactctggtg tcacaataaa 1680  
 cctgtatgta gctgtg 1696

<210> 3  
 <211> 260  
 <212> DNA  
 <213> *Rattus norvegicus*

<400> 3  
 ttctacatgg aggagggggc tccctactgc gagcgagatt acgagaagat gtttggcaca 60  
 aagtgtcgcg gctgtgactt caagatcgat gccggggacc gtttcctgga agcctgggt 120  
 ttcagctggc atgatacgtg ttttgtttgc gcaatatgtc aaatcaactt ggaaggaaag 180  
 accttctact ccaagaagga caagccctg tgcaagagcc atgccttttc ccacgtatga 240  
 gcacctctc acactactgc 260

<210> 4  
 <211> 16  
 <212> DNA  
 <213> MMLV

&lt;400&gt; 4

aagctttttt tttttg

16

&lt;210&gt; 5

&lt;211&gt; 13

&lt;212&gt; DNA

&lt;213&gt; MMLV

&lt;400&gt; 5

aagcttggct atg

13

&lt;210&gt; 6

&lt;211&gt; 223

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

atccttgctc	acctcacggg	caccgagttc	atgcaagacc	eggatgagga	gcacctgaag	60
aaatcaagcc	aggtgcccag	gacagaagcc	ccagccccag	cctcatctac	accccaggag	120
ccctggcctg	gccctaccgc	cccagccct	accagccgc	cgccctgggc	tgtggacct	180
gcgtttgccg	agcgtatgc	cccagacaaa	accagcacag	tgc		223

&lt;210&gt; 7

&lt;211&gt; 717

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 7

atggattcct	tcaaggtagt	gctggagggg	ccagcacctt	ggggcttccg	gctgcaaggg	60
ggcaaggact	tcaatgtgcc	cctctccatt	tcccggctca	ctcctggggg	caaagcggcg	120
caggccggag	tggccgtggg	tgactgggtg	ctgagcatcg	atggcgagaa	tgccggtagc	180
ctcacacaca	tgaagctca	gaacaagatc	cgggcctgcg	gggagcgctt	cagcctgggc	240
ctcagcaggg	cccagccggt	tcagagcaaa	ccgcagaagg	cctccgcccc	cgcgcggac	300
cctccgcggt	acacctttgc	acccagcgtc	tccctcaaca	agacggcccc	gccctttggg	360
gcgcccccg	ccgctgacag	cgccccgcaa	cagaatggac	agccgctccg	accgctggtc	420
ccagatgcca	gcaagcagcg	gctgatggag	aacacagagg	actggcgggc	gcggccgggg	480
acaggccagt	cgcgttcctt	ccgcatacct	gcccacctca	caggcaccga	gttcatgcaa	540
gacccggatg	aggagcacct	gaagaaatca	agccaggtgc	ccaggacaga	agccccagcc	600
ccagcctcat	ctacaccca	ggagccctgg	cctggcccta	ccgccccag	ccctaccagc	660
cgccccccct	gggctgtgga	ccctgcgttt	gccgagcgct	atgccccgga	caaaacg	717

&lt;210&gt; 8

&lt;211&gt; 1488

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 8

atcgatggcg	agaatgcggg	tagcctcaca	cacatcgaag	ctcagaacaa	gatccggggc	60
tgccgggagc	gcctcagcct	gggcctcagc	agggcccagc	cggttcagag	caaaccgcag	120
aaggcctccg	ccccgcgcgc	ggaccctccg	cggtagacct	ttgcacccag	cgtctccctc	180
aacaagacgg	cccgccctt	tggggcgccc	ccgcccgtg	acagcgcccc	gcaacagaat	240
ggacagccgc	tccgaccgct	ggtcccagat	gccagcaagc	agcggctgat	ggagaacaca	300

```

gaggactggc ggccgcggcc ggggacaggc cagtcgcgtt ccttccgcat ccttgccac 360
ctcacaggca ccgagttcat gcaagaccgc gatgaggagc acctgaagaa atcaagccag 420
gtgccaggga cagaagcccc agccccagcc tcatctacac cccaggagcc ctggcctggc 480
cctaccgccc ccagccctac cagccgcccg ccctgagctg tggaccctgc gtttgccgag 540
cgctatgccc cggacaaaac gagcacagtg ctgaccgggc acagccagcc ggccacgccc 600
acgccgctgc agagccgcac ctccattgtg caggcagctg ccggaggggg gccaggaggg 660
ggcagcaaca acggcaagac tcccgtgtgt caccagtgc acaagggtcat ccggggccgc 720
tacctgggtg cgttgggcca cgcgtaccac ccggaggagt ttgtgtgtag ccagtgtggg 780
aaggctcctg aagagggtg cttctttgag gagaaggggc ccatcttctg cccaccatgc 840
tatgacgtgc gctatgcacc cagctgtgcc aagtgcaga agaagattac aggcgagatc 900
atgcacgccc tgaagatgac ctggcacgtg cactgcttta cctgtgctgc ctgcaagacg 960
cccatccgga acagggcctt ctacatggag gagggcgtgc cctattgcga gcgagactat 1020
gagaagatgt ttggcacgaa atgccatggc tgtgacttca agatcgacgc tggggaccgc 1080
ttcctggagg ccctgggctt cagctggcat gacacctgct tcgtctgtgc gatatgtcag 1140
atcaacctgg aaggaaagac cttctactcc aagaaggaca ggctctctg caagagccat 1200
gccttctctc atgtgtgagc cccttctgcc cacagctgcc gcggtggccc ctagcctgag 1260
gggcctggag tcgtggccct gcatttctgg gtagggtggt caatgggtgc cttaacctg 1320
gtcctgggcc cgagcctggg ctcccgggcc cctgcccacc caccttatcc tcccaccca 1380
ctccctccac caccacagca caccggtgct ggccacacca gcccccttcc acctccagt 1440
ccacaataaa cctgtaccca gctgaattcc aaaaaatcca aaaaaaaa 1488

```

&lt;210&gt; 9

&lt;211&gt; 1644

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

```

atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaaggg 60
ggcaaggact tcaatgtgcc cctctccatt tcccggctca ctccctggggg caaagcgggc 120
caggccggag tggccgtggg tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 180
ctcacacaca tcgaagctca gaacaagatc cgggcctgcg gggagcgcct cagcctgggc 240
ctcagcaggg cccagccggt tcagagcaaa ccgcagaagg cctccgcccc cgcccgggac 300
cctccgcggt acacctttgc acccagcgct tccctcaaca agacggcccc gccctttggg 360
gcgccccgcg ccgctgacag cgcgccgcaa cagaatggac agccgctccg accgctggtc 420
ccagatggcca gcaagcagcg gctgatggag aacacagagg actggcgggc gcggccgggg 480
acaggccagt cgcgttcctt ccgcatcctt gccacacctc caggcaccga gttcatgcaa 540
gaccgggatg aggagcacct gaagaaatca agccaggtgc ccaggacaga agccccagcc 600
ccagcctcat ctacacccca ggagccctgg cctggcccta ccgccccag ccctaccagc 660
cgccccccct gggtgtgga cctgcgttt gccgagcgct atgccccgga caaaacgagc 720
acagtgtgta cccggcacag ccagccggcc acgcccacgc cgctgcagag ccgcacctcc 780
attgtgcagg cagctgccgg aggggtgcca ggagggggca gcaacaacgg caagactccc 840
gtgtgtcacc agtgccacaa ggtcatccgg gcccgctacc tgggtggcgtt gggccacgcg 900
taccacccgg aggagtttgt gtgtagccag tgtgggaagg tcctggaaga ggggtggctt 960
tttgaggaga agggcgccat cttctgcccc ccatgctatg acgtgcgcta tgcaccagc 1020
tgtgccaagt gcaagaagaa gattacaggg gagatcatgc acgcccctga gatgacctg 1080
cacgtgcact gctttacctg tgctgcctgc aagacgcccc tccggaacag ggccttctac 1140
atggaggagg gcgtgcccta ttgcgagcga gactatgaga agatgtttgg cacgaaatgc 1200
catggctgtg acttcaagat cgacgctggg gaccgcttcc tggaggccct gggttccagc 1260
tggcatgaca cctgcttcgt ctgtgcgata tgtcagatca acctggaagg aaagaccttc 1320
tactccaaga aggacaggcc tctctgcaag agccatgcct tctctcatgt gtgagccct 1380
tctgcccaca gctgcgcggg tggcccttag cctgaagggc ctggagtctg ggcctgtcat 1440
ttctgggtag ggctggcaat ggttgcccta accctggctc ctggcccgag cctgggctcc 1500
cgggccctg cccaccacc ttatcctccc accccactcc ctccaccacc acagcacacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgcac aataaacctg taccagctg 1620
aattccaaaa aatccaaaa aaaa 1644

```

&lt;210&gt; 10

&lt;211&gt; 457

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe  
 1 5 10 15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg  
 20 25 30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp  
 35 40 45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile  
 50 55 60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly  
 65 70 75 80

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala  
 85 90 95

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu  
 100 105 110

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala  
 115 120 125

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser  
 130 135 140

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly  
 145 150 155 160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln  
 180 185 190

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu  
 195 200 205

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220

Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240

Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255

Ser Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly



260	265	270
Gly Ser Asn Asn Gly Lys Thr	Pro Val Cys His Gln Cys His Lys Val	
275	280	285
Ile Arg Gly Arg Tyr Leu Val	Ala Leu Gly His Ala Tyr His Pro Glu	
290	295	300
Glu Phe Val Cys Ser Gln Cys Gly Lys Val	Leu Glu Glu Gly Gly Phe	
305	310	315
Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro	Pro Cys Tyr Asp Val Arg	
325	330	335
Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys	Lys Ile Thr Gly Glu Ile	
340	345	350
Met His Ala Leu Lys Met Thr Trp His Val	His Cys Phe Thr Cys Ala	
355	360	365
Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala	Phe Tyr Met Glu Glu Gly	
370	375	380
Val Pro Tyr Cys Glu Arg Asp Tyr Glu Lys	Met Phe Gly Thr Lys Cys	
385	390	395
His Gly Cys Asp Phe Lys Ile Asp Ala Gly	Asp Arg Phe Leu Glu Ala	
405	410	415
Leu Gly Phe Ser Trp His Asp Thr Cys Phe	Val Cys Ala Ile Cys Gln	
420	425	430
Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser	Lys Lys Asp Arg Pro Leu	
435	440	445
Cys Lys Ser His Ala Phe Ser His Val		
450	455	

&lt;210&gt; 11

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 11

gccagggttt tcccagtcac ga

22

&lt;210&gt; 12

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 12

gccagggttt tcccagtcac ga

22

<210> 13  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 13  
tcttagcaga gccagcctg ct 22

<210> 14  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 14  
gcatgaactc tgtgcccgtg ag 22

<210> 15  
<211> 20  
<212> DNA  
<213> Rattus norvegicus

<400> 15  
atccttgctc acctcacggg 20

<210> 16  
<211> 22  
<212> DNA  
<213> Rattus norvegicus

<400> 16  
gcactgtgct ggttttgtct gg 22

<210> 17  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 17  
catggattcc ttcaaggtag tgc 23

<210> 18  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 18  
gttttgtctg gggcagagcg 20

<210> 19  
<211> 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: adaptor for  
Marathon RACE reactions

&lt;400&gt; 19

ctaatacgac tcactatagg gctcgagcgg ccgcccgggc aggt

44

&lt;210&gt; 20

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: PCR primer  
specific for Marathon RACE adaptor

&lt;400&gt; 20

ccatccta acgactcact atagggc

27

&lt;210&gt; 21

&lt;211&gt; 765

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

```

ccgttggttg taaaacgacg cagagcagcg ccctggccgg gccaaagcagg agccggcattc 60
atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaagg 120
ggcaaggact tcaatgtgcc ctctccatt tcccggtca cctctggggg caaggccgtg 180
caggccggag tggccgtaag tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 240
ctcacacaca tcgaagctca gaacaagatc cgggcctgcg gggagcgcct cagcctgggc 300
ctcaacaggg cccagccggt tcagaacaaa ccgcaaaagg cctccgcccc cgccgcggac 360
cctccgcggg acacctttgc accaagcgtc tccctcaaca agacggcccc gcccttgggg 420
gcgccccgcg ccgctgacag cgccccgcag cagaatggac agccgctccg accgctggc 480
ccagatgccg gcaagcagcg gctgatggag aacacagagg actggcgggc gcggccgggg 540
acaggccagt gccgttcctt tcgcatcctt gctcacctta caggcaccga gttcatgcaa 600
gacccggtat aggagcacct gaagaaatca agccaggtgc ccaggacaga agccccagcc 660
ccagcctcat ctacaccca ggagccctgg cctggcccta ccgccccag ccctaccagc 720
cgccccccct gggctgtgga ccctgcgttt gccgagcgct atgcc 765

```

&lt;210&gt; 22

&lt;211&gt; 1689

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

```

cgacgcagag cagcgccctg gccgggcca gaggagccg gcatcatgga ttccttcaag 60
gtagtgtggt aggggcccag accttggggc tccggctgc aagggggcaa ggacttcaat 120
gtgcccctct ccatttcccg gctcactcct gggggcaaa cggcgaggc cggagtggcc 180
gtgggtgact ggggtgctgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240
gtcagaaca agatccgggc ctgccccggg cgcctcagcc tgggcctcag cagggccag 300

```

```

ccggttcaga gcaaaccgca gaaggcctcc gccccgcgcg cggaccctcc gcggtacacc 360
tttgcaccca gcgtctccct caacaagacg gcccggccct ttggggcgcc cccgcccgt 420
gacagcgccc cgcaacagaa tggacagccg ctccgaccgc tggccccaga tgccagcaag 480
cagcggtga tggagaacac agaggactgg cggccgcggc cggggacagg ccagtcgcgt 540
tccttccgca tccttgccca cctcacaggc accgagttca tgcaagaccc ggatgaggag 600
cacctgaaga aatcaagcca ggtgcccagg acagaagccc cagccccagc ctcatctaca 660
cccaggagc cctggcctgg ccctaccgcc ccagcccta ccagccgcc gccctgggct 720
gtggaccctg cgtttgccga gcgctatgcc ccggacaaaa cgagcacagt gctgaccgg 780
cacagccagc cggccacgcc cagcccgctg cagagccgca cctccattgt gcaggcagct 840
gccggagggg tgccaggagg gggcagcaac aacggcaaga ctcccgtgtg tcaccagtgc 900
cacaagggtca tccggggcgc ctacctggtg gcgttgggccc acgcgtacca cccggaggag 960
tttgtgtgta gccagtgtgg gaaggctcctg gaagagggtg gcttctttga ggagaagggc 1020
gccatcttct gccaccatg ctatgacgtg cgctatgcac ccagctgtgc caagtgcag 1080
aagaagatta caggcgagat catgcacgcc ctgaagatga cctggcacgt gcactgcttt 1140
acctgtgctg cctgcaagac gcccatccgg aacagggcct tctacatgga ggagggcggtg 1200
ccctattgctg agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagatcgacg ctggggaccg ctccctggag gccctgggct tcagctggca tgacacctgc 1320
ttcgtctgtg cgatatgtca gatcaacctg gaaggaaaga ccttctactc caagaaggac 1380
aggcctctct gcaagagcca tgccttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcctgga gtogtggccc tgcatttctg ggtagggctg 1500
gcaatgggtg ccttaacctt ggctcctggc ccgagcctgg gctcccgggc ccctgcccac 1560
ccaccttacc ctcccacccc actccctcca ccaccacagc acaccgggtg tggccacacc 1620
agcccccttt cacctccagt gccacaataa acctgtacct agctgaattc caaaaaatcc 1680
aaaaaaaaa

```

<210> 23  
 <211> 22  
 <212> DNA  
 <213> Homo sapiens

<400> 23  
 gcactgtgct cgttttgtcc gg 22

<210> 24  
 <211> 21  
 <212> DNA  
 <213> Homo sapiens

<400> 24  
 tccttgctca cctcacgggc a 21

<210> 25  
 <211> 30  
 <212> DNA  
 <213> Homo sapiens

<400> 25  
 tcctcatccg ggtcttgcac gaactcgggtg 30

<210> 26  
 <211> 28  
 <212> DNA

<213> Homo sapiens

<400> 26

gcccccgccc gctgacagcg ccccgcaa

28

<210> 27

<211> 24

<212> DNA

<213> Homo sapiens

<400> 27

tccttgctca cctcacgggc accg

24

<210> 28

<211> 22

<212> DNA

<213> Homo sapiens

<400> 28

gtaatacgac tcactatagg gc

22

<210> 29

<211> 23

<212> DNA

<213> Rattus norvegicus

<400> 29

gcggctgatg gagaatactg aag

23

<210> 30

<211> 23

<212> DNA

<213> Rattus norvegicus

<400> 30

atcttggtggc actggtggca tac

23

<210> 31

<211> 22

<212> DNA

<213> Rattus norvegicus

<400> 31

tgtgtcgggt cagcactgtg ct

22

<210> 32

<211> 1620

<212> DNA

<213> Homo sapiens

&lt;400&gt; 32

```

atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaaggg 60
ggcaaggact tcaatgtgcc cctctccatt tcccggctca ctctggggg caaagcggcg 120
caggccggag tggccgtggg tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 180
ctcacacaca tcgaagctca gaacaagatc cgggcctgcg gggagcgctt cagcctgggc 240
ctcagcaggg cccagccggt tcagagcaaa ccgcagaagg cctccgcccc cgccgcgga 300
cctcccggtt acacctttgc acccagcgtc tccctcaaca agacggcccc gccctttggg 360
gcgcccccg cgcgtgacag cgccccgcaa cagaatggac agccgctccg accgctggtc 420
ccagatgccg gcaagcagcg gctgatggag aacacagagg actggcgggc gcggccgggg 480
acaggccagt cgcgttccct ccgcctcctt gcccacctca caggcaccca gttcatgcaa 540
gacccggatg aggagcacct gaagaaatca agccaggtgc ccaggacaga agccccagcc 600
ccagcctcat ctacacccca ggagccctgg cctggcccta ccgccccag ccctaccagc 660
cgccccctt gagctgtgga ccctgcgttt gccgagcgct atgccccgga caaacgagc 720
acagtgtgta cccggcacag ccagccggcc acgcccacgc cgctgcagag ccgcacctcc 780
attgtgcagg cagctgccgg aggggtgcca ggagggggca gcaacaacgg caagactccc 840
gtgtgtcacc agtgccacaa ggtcatccgg ggccgctacc tgggtggcgtt gggccacgcg 900
taccacccgg aggagtttgt gtgtagccag tgtgggaagg tcctggaaga ggggtgcttc 960
tttgaggaga agggcgccat cttctgcca ccatgctatg acgtgcgcta tgcaccagc 1020
tgtgccaagt gcaagaagaa gattacaggc gagatcatgc acgccctgaa gatgacctgg 1080
caegtgcact gctttacctg tgctgcctgc aagacgcccc tccggaacag gcccttctac 1140
atggaggagg cgtgccccta ttgcgagcga gactatgaga agatgtttgg cacgaaatgc 1200
catggctgtg acttcaagat cgacgctggg gaccgcttcc tggaggccct gggcttcagg 1260
tggtcatgaca cctgcttcgt ctgtgcgata tgtcagatca acctggaagg aaagaccttc 1320
tactccaaga aggacaggcc tctctgcaag agccatgcct tctctcatgt gtgagcccc 1380
tctgcccaca gctgcgcggg tggcccctag cctgaggggc ctggagtcgt ggccctgcat 1440
ttctgggtag ggctggcaat ggttgcccta acctggctc ctggcccag cctgggctcc 1500
cgggcccctg cccacccacc ttatcctccc acccactcc ctccaccacc acagcacacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgccac aataaacctg taccagctg 1620

```

&lt;210&gt; 33

&lt;211&gt; 1665

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 33

```

cgacgcagag cagcgccctg gccggggccaa gcaggagccg gcatcatgga ttcttcaag 60
gtagtgtggt agggggccagc accttggggc tccgggtgc aagggggcaa ggacttcaat 120
gtgcccctct ccatttcccg gctcactcct gggggcacaag cggcgagggc cggagtggcc 180
gtgggtgact ggggtgtgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240
gctcagaaca agatccgggc ctgcggggag cgcctcagcc tgggcctcag cagggccag 300
ccggttcaga gcaaacgcga gaaggcctcc gccccgcggc cggaccctcc gcgtacacc 360
tttgacccca gcgtctccct caacaagacg gcccgggcct tggggcgccc ccgcccgt 420
gacagcgccc cgcaacagaa tggacagccg ctccgaccgc tgggtccaga tgccagcaag 480
cagcggtgta tggagaacac agaggactgg cggccgcggc cggggacagg ccagtgcgct 540
tccttccgca tccttgccca cctcacaggc accgagttca tgcaagacct ggatgaggag 600
cacctgaaga aatcaagcca ggtgcccagg acagaagccc cagccccagc ctcatctaca 660
ccccaggagc cctggcctgg ccctaccgcc ccagcccta ccagccgcc gccctgagct 720
gtggaccctg cgtttgccga gcgctatgcc ccggacaaaa cgagcacagt gctgaccgg 780
cacagccagc cggccacgcc cagccgctg cagagccgca cctccattgt gcaggcagct 840
gccggagggg tgccaggagg gggcagcaac aacggcaaga ctcccgtgtg tcaccagtgc 900
cacaaggtca tccggggcgc ctacctggtg gcgttgggac acgctacca cccggaggag 960
tttgtgtgta gccagtgtgg gaaggtcctg gaagagggtg gcttctttga ggagaagggc 1020
gccatcttct gccccacctg ctatgacgtg cgctatgcac ccagctgtgc caagtgcag 1080
aagaagatta caggcgagat catgcacgcc ctgaagatga cctggcacgt gcactgcttt 1140
acctgtgctg cctgcaagac gcccatccgg aacagggcct tctacatgga ggagggcggtg 1200

```

```

ccctattgcg agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagatcgacg ctggggaccg cttcctggag gccctgggct tcagctggca tgacacctgc 1320
ttcgtctgtg cgatatgtca gatcaacctg gaaggaaaga ctttctactc caagaaggac 1380
aggcctctct gcaagagcca tgccttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcctgga gtcgtggccc tgcatttctg ggtagggetg 1500
gcaatgggtg ccttaaccct ggctcctggc ccgagcctgg gctcccgggc ccctgcccac 1560
ccaccttatc ctcccacccc actccctcca ccaccacagc acaccggtgc tggccacacc 1620
agcccccttt cacctccagt gccacaataa acctgtaccc agctg 1665

```

&lt;210&gt; 34

&lt;211&gt; 223

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 34

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1             5             10             15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Scr Arg
      20             25             30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
      35             40             45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
      50             55             60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
      65             70             75             80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
      85             90             95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
      100            105            110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
      115            120            125

```

```

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser
      130            135            140

```

```

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly
      145            150            155            160

```

```

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr
      165            170            175

```

```

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln
      180            185            190

```

```

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu
      195            200            205

```

```

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro

```

210

215

220

<210> 35  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 35  
 gagccggcat catggattcc

20

<210> 36  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 gctgcctgca caatggaggt

20

<210> 37  
 <211> 1456  
 <212> DNA  
 <213> Homo sapiens

<400> 37  
 cgacgcagag cagcgccctg gccggggccaa gcaggagccg gcatcatgga ttccttcaag 60  
 gtagtgctgg agggggccagc accttggggc ttccggctgc aagggggcaa ggacttcaat 120  
 gtgccccctc ccatttcccg gctcactcct gggggcaaaag cggcgaggc cggagtggcc 180  
 gtgggtgact ggggtgctgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240  
 gctcagaaca agatccgggc ctgcggggag cgcctcagcc tgggcctcag cagggcccag 300  
 ccggttcaga gcaaaccgca gaaggtgcag acccctgaca aacagccgct ccgaccgctg 360  
 gtcccagatg ccagcaagca ggggctgatg gagaacacag aggactggcg gccgcggccg 420  
 gggacaggcc agtcgcgttc ctcccgcatc cttgcccacc tcacaggcac cgagttcatg 480  
 caagaccggg atgaggagca cctgaagaaa tcaagccagg tgcccaggac agaagcccca 540  
 gccccagcct catctacacc ccaggagccc tggcctggcc ctaccgcccc cagccctacc 600  
 agccgcccgc cctgggctgt ggaccctgcg tttgccgagc gctatgcccc ggacaaaacg 660  
 agcacagtgc tgaccgggca cagccagccg gccacgccc cgcgctgca gagccgcacc 720  
 tccattgtgc aggcagctgc cggaggggtg ccaggagggg gcagcaacaa cggcaagact 780  
 cccgtgtgtc accagtgccca caaggtcatc cggggccgct acctggtggc gttggggccac 840  
 gcgtaccacc cggaggagtt tgtgtgtagc cagtgtggga aggtcctgga agagggtggc 900  
 ttctttgagg agaagggcgc catcttctgc ccaccatgct atgacgtgcg ctatgcaccc 960  
 agctgtgccca agtgcaagaa gaagattaca ggcgagatca tgcacgccct gaagatgacc 1020  
 tggcacgtgc actgctttac ctgtgctgcc tgcaagacgc ccatccggaa cagggccttc 1080  
 tacatggagg agggcgtgcc ctattgcgag cgagactatg agaagatgtt tggcacgaaa 1140  
 tgccatggct gtgacttcaa gatcgacgct ggggaccgct tcctggaggc cctgggcttc 1200  
 agctggcatg acacctgctt cgtctgtgcg atatgtcaga tcaacctgga aggaaagacc 1260  
 ttctactcca agaaggacag gcctctctgc aagagccatg cttctctca tgtgtgagcc 1320  
 ccttctgccc acagctgccg cgggtggccc tagcctgagg ggctggagt cgtggccctg 1380  
 catttctggg tagggctggc aatggttgcc ttaaccctgg ctctggccc gagcctgggc 1440  
 tcccgggccc tgccc 1456

<210> 38  
 <211> 423



&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 38

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1              5              10              15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
          20              25              30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
          35              40              45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
          50              55              60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
          65              70              75              80

Leu Ser Arg Ala Gln Pro Val Gln Asn Lys Pro Gln Lys Val Gln Thr
          85              90              95

Pro Asp Lys Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser Lys Gln
          100             105             110

Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly Thr Gly
          115             120             125

Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr Glu Phe
          130             135             140

Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln Val Pro
          145             150             155             160

Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu Pro Trp
          165             170             175

Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp Ala Val
          180             185             190

Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser Thr Val
          195             200             205

Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln Ser Arg
          210             215             220

Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly Gly Ser
          225             230             235             240

Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Gln Val Ile Arg
          245             250             255

Ala Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu Glu Phe
          260             265             270

Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe Phe Glu

```

275	280	285
Glu Lys Gly Ala Ile Phe Cys Pro Pro Cys Tyr Asp Val Arg Tyr Ala		
290	295	300
Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile Met His		
305	310	315
Ala Leu Lys Met Thr Trp His Val Leu Cys Phe Thr Cys Ala Ala Cys		
325	330	335
Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly Val Pro		
340	345	350
Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys Gln Trp		
355	360	365
Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala Leu Gly		
370	375	380
Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln Ile Asn		
385	390	395
Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Arg Pro Leu Cys Lys		
405	410	415
Ser His Ala Phe Ser His Val		
420		

&lt;210&gt; 39

&lt;211&gt; 1575

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

```

cgacgcagag cagcgccctg gccggggccaa gcaggagccg gcatcatgga ttccttcaag 60
gtagtgtctgg agggggccagc accttggggc ttccggctgc aagggggcaa ggacttcaat 120
gtgcccctct ccatttcccg gctcactcct gggggcaaag cggcgcaggc cggagtggcc 180
gtgggtgact ggggtgtgag catcgatggc gagaatgcgg gtagcctcac acacatcgaa 240
gctcagaaca agatccgggc ctgcggggag cgcctcagcc tgggcctcag cagggccag 300
ccggttcaga gcaaaccgca gaaggcctcc gccccgccc cggaccctcc gcgttacacc 360
tttgcaccca gcgtctccct caacaagacg gcccggccct ttggggcgcc cccgccgct 420
gacagcgccc cgcaacagaa tgggtgcaga cccctgacaa acagccgctc cgaccgctgg 480
tcccagatgc cagcaagcag cggctgatgg agaacacaga ggactggcgg ccgcggccgg 540
ggacaggcca gtcgcgttcc ttccgcatcc ttgcccacct cacaggcacc gagttcatgc 600
aagaccgga tgaggagcac ctgaagaaat caagccaggt gccaggaca gaagccccag 660
ccccagctc atctacaccc caggagccct ggcctggccc taccgcccc agcctacca 720
gccgcccgc ctgggctgtg gaccctgcgt ttgccgagcg ctatgcccc gacaaaacga 780
gcacagtgtc gaccggcac agccagccgg ccacgcccac gccgctgcag agccgacct 840
ccattgtgca ggcagctgcc ggaggggtgc caggaggggg cagcaacaac ggcaagactc 900
ccgtgtgtca ccagtgcac aaggtcatcc ggggcccgtc cctggtggcg ttgggccag 960
cgtaccaccc ggaggagttt gtgtgtagcc agtgtgggaa ggtcctggaa gaggtggct 1020
tctttgagga gaagggcgcc atcttctgcc caccatgcta tgacgtgcgc tatgcacca 1080
gctgtgcaa gtgcaagaag aagattacag gcgagatcat gcacgccctg aagatgacct 1140
ggcacgtgca ctgctttacc tgtgctgcct gcaagacgcc catccggaac agggccttct 1200

```

```

acatggagga gggcgtgccc tattgcgagc gagactatga gaagatgttt ggcacgaaat 1260
gccatggctg tgacttcaag atcgacgctg gggaccgctt cctggaggcc ctgggcttca 1320
gctggcatga cacctgcttc gtctgtgcga tatgtcagat caacctggaa ggaaagacct 1380
tctactccaa gaaggacagg cctctctgca agagccatgc cttctctcat gtgtgagccc 1440
cttctgcccc cagctgccgc ggtggccctt agcctgaggg gcctggagtc gtggccctgc 1500
atttctgggt agggctggca atggttgect taacctggc tcctggccc agcctgggct 1560
cccgggccct gcca                                     1575

```

&lt;210&gt; 40

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1              5              10              15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
          20              25              30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
  35              40              45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
  50              55              60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
  65              70              75              80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
          85              90              95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
  100              105              110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
  115              120              125

```

```

Pro Gln Gln Asn Gly Cys Arg Pro Leu Thr Asn Ser Arg Ser Asp Arg
  130              135              140

```

```

Trp Ser Gln Met Pro Ala Ser Ser Gly
  145              150

```

&lt;210&gt; 41

&lt;211&gt; 24740

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 41

```

nnnnnntgta ttttatcata ttttaaaaat caaaaaacaa aaggcagttg aggttaggca 60
tggagggttc tgctgtaat cccagcactt tgggaagccg aagcacgtgg atcacctgag 120
gtcaggagtt cgagaccagc ctgcccaata tggtaaaacc ctgtctctac taaaaatata 180
aaaaattagc caggcatggt ggtgggcacc tgtaatccca gctacttggg agactgaggc 240

```

```

aggagaatca cttaaacccg ggaggcgggc tgggcgcggg ggctcatgcc tgtaatccca 300
gcactttggg aggccgagac aggcggatca tgaggtcagg agatcgagat catcctggct 360
aacatggtga aaccccatct ctactaaaaa tacaaaaaaa attagccagg cctgggtggcg 420
ggcacctgta gtcccagcta cttgggaggc tgaggcagga gaatggcgtg aacctgggag 480
gcggcgttgc agtgagccaa gatcgcgcca ctgcactcca gcctgggcga caagagttag 540
actccatctt aaagaaaaaa aacaaacccg ggaggcggaa attgcagtca gccgagatct 600
cgccattgca ctcaagtatg ggtgacagag caagactcca tgtcaaaaaa aaaggcagtt 660
gacaggagca aggagcctgg tgaggaagct gtggcatttg acccggtgtg gttgctatgg 720
gccaggggtg tgctagtaga ggagctgagt gggaaagagc acaggggaca tgctgaaggc 780
ctgggtgtgg ggatgaggca gagattgggg gcaccttgca gggtcatagc aggtggctgt 840
ggtgagatgg aggaagacac ctgggggtact gctctaggct gtcagacata cagaagctgg 900
cccagccaag cccaggggct gcaagggaca tccttttgtg tccccagtga tctgcagctc 960
tcagacaccc tcaagcacag tgccctctgc ccagcccagc actctcagtg gggagccagg 1020
tgggagaaca ggctcggaag gggacctagg cttatgcagc gagccgggca aagctggaac 1080
tggagcccag gcccctggat gccccctggc ttgtggagtt ctgggatact gaggggaggg 1140
gacagggcat gggagtgcgg tgcctcacc tttgacttga actcattccc caggggacag 1200
gggaggcctc ctcaggatcc acagatgcc agtctcccaa gaggggcctg gtccccatgg 1260
aggaaaactc catctactcc tcctggcagg aaggttaagt ggaggacgtg caagggcagc 1320
ctcagccccc cacaccagg gctgggtctt tttgggactg acggagctgt cctggccacc 1380
tgccacagtg ggcgagtttc ccgtgggtgg gacaggact gagggcgcca cccgtgcca 1440
gtggaagggg ccggccctgc tgggtgctgg ccagacgcc atccagctga gggagtgcaa 1500
ggcaccaggg ccctctacag ctggccctac cacttcctgc gcaagtctgg ctccgacaag 1560
gtgaggtgca ggggtgggaa agggtgaggg gctgacagcc tggaccctcc tgctaattcc 1620
caccctgtgt ccctgtgccc agggcgtgtt ctcctttgag gccggccgtc gctgccactc 1680
gggtgagggc ctctttgcct tcagcacccc ctgtgcccct gacctgtgca gggctgtggc 1740
cggggccatc gccgccagcg ggagcggctg ccagagctga ccaggcccca gccctgcccc 1800
ctgccacggg ccacctctct gccctccctg gacacccccg gagagcttcg ggagatgcca 1860
ccaggacctg agccaccac gtccaggaaa atgcacctgg ccgagcccg accccagagc 1920
ctgccgctac tgctaggccc ggagcccaac gatctggcgt ccgggctcta cgcttcagtg 1980
tgcaagcgtg ccagtgggcc cccaggcaat gagcacctct atgagaacct gtgtgtgctg 2040
gaggccagcc ccacgtgca cgggtgggaa cctgagccgc acgagggccc cggcagccgc 2100
agccccacaa ccagtcccat ctaccacaac ggccaggact tgagctggcc cggccccggc 2160
aacgacagta ccctggaggc ccagtaccgg cggctgctgg agctggatca ggtggagggc 2220
acaggccgcc ctgacctca ggcaggttcc aaggccaaagc tggtagccct gctgagtcgt 2280
gagcggagga agggccagc cccttgtagc cgccctgaa cggccagcag agtggtgccc 2340
agaggggaga ggtgctcccc ctgggacagg aggggtgggt ggtgggcaaa cattggcccc 2400
atgcagacac acgcctgtgt ccacctggc ctgcaggaa aaggcaggcc gcctgtggag 2460
gacctcagcc ctgccctgcc ctcctcatga atagtgtgca gactcacaga taataaagct 2520
cagagcagct cccggcaggg gcactcacgg cacacgcccc tgcccacgtt cattgcccgc 2580
aacacaagca ccctgtgccg gttccagggg cacaggtgac ctgggcctta cctgccaccc 2640
gtgggctcaa acccactgca gcagacagac gggatggaaa tcattaggac tccatgttgc 2700
tctgcacggc cgagtgcac gaagaggcag gcggaggag ctgtgaggct tacttgtcag 2760
actcaggaag gagcaacatg agggcccaac tggagacccg gagggccgag ctgggaggag 2820
gcagtggggg cgggtgtag gtggaaggga tttcagagac accctcgtcc aaaacacttg 2880
ttccctgctg aaactccaac aatttgtaga tacttctggg aaccccaggc gtcagtctcc 2940
tcatctgtaa aggagagaga accgatgacg tatcaggcat aatccttgat gagagtttgc 3000
tgctgtgcta ctcagtgcga ggcgtgggg gacacagccg tgctcaggac agccttggtc 3060
ctgttctccg ggagccgaca ttccaggggg agagaagttt cctgaagact tccatgctgc 3120
gttccctcct ctgctcctgc tcctggcgcc atcctaggag ccagccatgc acgtaagcgt 3180
catgcctcca gggctctgac tgcccagccc ctcaccgcaa ctccacctca gctgcacaca 3240
cccttggcac atcctgaacc tcattttcat gacggacaca caatttttgc tctctcctgt 3300
ccaagcctca tcctctggcc gccacctcct tccagctcac ttcctttagt gcgccagta 3360
ccgccccctg ctaggcatgt cgacctgcag ggacctttt ctggctcttc gaggcctctg 3420
cccaccatcc cctctttgtt ctccatagtc ccttccccct gttctctctc gtttcatctt 3480
actggtctgg caaagtcccc ggccttgggc gagccagacc tcctcagtgc ctgcacacag 3540
ctgcccacag ccagagaaat ccatttaagc agactgcctg catccttctt aacagtgcaa 3600

```

```

ggcaggcact ccctgccaca agagaccctg ttccctagta gggcagcttt tctcctcccc 3660
agaacctcct gtctatcccc acccaatgtc tcctcacagg catattgggg aaacagggtca 3720
ggctctccca cgtatctctg aagtgtactg gcatccatct gtcttcttcc tacccttaca 3780
gtagaaacag tgtctgtccc cagctgtgct ctgatcccgg ctcttttcac ctcagagctt 3840
ggaaaattga gctgtcccca ctctctcctg cgccattca tcctaccagc agcttttcca 3900
gccacacgca aacatgtctt gtaatttcac attttaaacc ttcccttgac ctcacattcc 3960
tcttcggcca cctctgtttc tctgttcttc ttcacagcaa aaactgttca aaagagtgtg 4020
tgattacttt catttccact ttctcacccc cattctctcc tcaattaact ctcttctatc 4080
cccagatgac cattatgtgg cttttattag agtcaccaac cttattctcc aaaacaaaag 4140
caacaaggac tttgacttct cagcagcact cagctctggg tcttgaaaca cccccgttac 4200
ttgctattcc tcctacctca taacaatctc ctcccagcc tctactgctg ctttctctga 4260
gttcttccca gggctcctagg ctcagatgta gtgtagctca accctgctac acaaagaatc 4320
tcctgaaagc ctgtaaaaaat gtccatgcat gttctgtgag tgatctacca agaaaataaa 4380
aaattttaaa aatcaaatgc ccagtcctgg gccacacgc aggggctctg atttcatcag 4440
tctggtaggt ggggttctggg catccacgct cactggattt ccggatgatt gtagtatgca 4500
gcctaggctg ggaacctact gcctcagcaa gccagtcatt ctccagggtg cacagaccct 4560
ctaggtgcta atgaccccgga aggtctgtct tcagtgcaca cctccccctg agctccagat 4620
ttaggaattcc cactgcacac gagacatctg gatgtggaaa agacatctcc agatcccatg 4680
ggtgaaaggg ggttggggga atggagactc gtgttcttcc aggatgtgtg tggacacaga 4740
atgcaaagcc tggaggggat ctagagccat agggaggaag atttcggctc acttattcat 4800
gcaagcactt cctgatgggt aaggcttag agcaagctga ggccaagagg cgggcagtcg 4860
aggtgctgct gcaggcacc ccactcccta cagtggcaag cccaagcca gcccttgga 4920
gctcaaattc caggacacgc tgaaggctcac ccagagagtc aggggcatgg ctagaaccag 4980
aaccaggac tctggggacc cagcatggca tcctttcctt cattacaaat ctgagctgct 5040
ttgtttccta gggatttctg tgatattcca aggggactgt gggaaagaaa gtccttgga 5100
accaccagga cgctagaggc ctggcctgga gcctcaggag tctcggccac cagagggcgc 5160
tgggtccttg tccaggctca gttgctacgc aggggctgcc tgtgctggga ggctcccag 5220
gggacacaga ccagagcctt gcaccagccc aaggaatggg agcctggggg cctctctgct 5280
ggaggactgc caggaccccc aggtgcccgc ctcttcttct gctcatttgc tgtttcactt 5340
tgtcaatcct tcctttcttc gtgtgttcat tcacatccac tgtgtgctgg ccctggggaa 5400
atgttagata agacacatta gctgtgtgtc ttcatgttcc taacaaagaa cacaccctgg 5460
aaagtagacc gcagagagtc cccattcccc cactccctc cacacatgga atctggagat 5520
gccttttcca catccagatg tctctggtgc tgtgggattc ttaaataaac aaacatttca 5580
tacagaatgt gagatgatgg agatgctatg gggaaaagta aagcagaggg agggcctagt 5640
gtgtgatgcg ggtgaggcat ccagggattg ctgtttcagc tgtgatcagg aaaggccctg 5700
ggaggaggcc acatctgagc agagacctaa ataaagtgtg aaacctgttg ctgagatatt 5760
tggagaagtg tttcaagggc cgggcaccgg gcatggtggc tcacgcctgt aatcccagca 5820
ctttgggagg ccaaggcagg tggatcgctg gaggtcagga gtttgagagc agcctgacca 5880
acatggagaa accccatctc tactaaacat ataaaaatta tccgggcatg gtggttcatg 5940
cctgtagtcc cagctactcg ggaggttgag gcaggagaat cacttgaacg tgggaggcag 6000
aggttgacgc aagccgagat cacaccactg cactccagcc tggatgacag agcagactc 6060
cgtctcaaaa aaaaaaaga aaagaaaaaa gaaaaaaa gaaaagtgtt tcaagcaggg 6120
gaactggcaa gtggagaggc cctgaggcag aaatatgctt ggctgctgg aggaaatgtg 6180
agtgaggagg tcagggtggc tggagtggag ggagcgagtg gtaggagtca gaccagttt 6240
attcatattc ttaggtctt aaggacttca gttttatttt gagtgcata tgagccact 6300
ggaatgctaa aagctgagag tgacatgggt ctgtgattct ggctttaaaa atatcacttt 6360
ggctgcttcg tgaagactct ggaaggggca agggtgaaag cagggatgcc cgttaggaga 6420
ccgttacagg ggcgcaggca caaaatggca gtggctggga caatggtggc agcagcgggt 6480
agatgtgaac atgttgaagg tggaaatttc agaactctgg ggaggacaga agagaaagga 6540
taacttctc gtttctgctg aaccagtgtg ataaatgttg gtggcacttc ttgaagttag 6600
gaaggagtta ggaagggtgg aaaggcacia gtttgaattg ggccatgatg gtctgagata 6660
cctagtacag tggttcccca acctttttgg cagaagggac cgctttcatg gaagacaatt 6720
tttccacaga ctgggggtgg ggtggggatg gtttcagggt ggctcgagtg cagtacattt 6780
atcattagac tctttttttt tttttttttt tgagatagag tctcgctctg tcaccacac 6840
tggagtgcag tggagccatc ttggctcact acaacctctg ctgcccagg tcaagtcat 6900
ctcctgcctc agcctctcaa gtagctggga ttataggcat atgcgccacc acgcccagct 6960

```

```

aattttttagta ttttttagtag agacgggggtt tcaccatatt ggccaggatg gtctcgaact 7020
cctgacctca agtgatectc cccgcctca acctcccaa gtgctgggt tacaggcgtg 7080
aaccactgca cccggcccat ttatcattag attctcataa ggaatgagca acctagatcc 7140
ctcgcatgca cagttcaca tagggttcac gctcctatgg gagtctaag ctgccgctgc 7200
actcagcttc tctggcttgc cgctgctcac cttctgctgt gcagcccagt tcctaacagg 7260
ccacaaacgg ggagttgggg acccctgac tagtaaaca ctaggcaggg ttttgataa 7320
tggagttaga gttcctgggg agaggtcagg ctggccatga aacatgggat gcctttgcat 7380
ataggtggtg ttgaaagcca caggacagta cggggtctca gggggtgagc ataaagagag 7440
gcgacatcag atggccaagg ccagaggcag aggaggatgg gaaggagggg ccagtggggc 7500
agggggaagc tgtgaagcca gggaaaaagg gtgtttcgcg gaaaaggatc aacctggacc 7560
agtgtgccc ctaggcaggg caggatgaaa ctttaaccacc acggattcca tggcccatg 7620
gcctccaggc cacaggggac cttgagaaga gagatctcag gggacgggtg cggacaagag 7680
ccgcctggc atggcttcaa gagataactg aaggaaagca agtgagagc cgataaacag 7740
acaactccct ggaggaattt tactctcgag aggagaatta aagggtagta gctggagagg 7800
gatgtgggg caagagaagg tctttaacga cgagaactct cacggcgggt tgtgcagaac 7860
aggggtgggtg tgatgactgt ggatggagag gggagaactg cagcgactct gtcctaggag 7920
gaggtgatgg gccgggacca ccaagcgagt ggaggggtga cggccctcc ctcacccgca 7980
caccgcagtg tgctcagtg cctgcccgc cgtggcctga cctgggctga acgaggggcc 8040
gggactctga ggacgcctcc caggcgcgca gtccgtctgg ccaagggtga gcgggacggc 8100
ngcttccgac ggtgcgggg tgggctcggg gttgcaggga catccggcgt ccgctcctgc 8160
cctgttttcc tgcttccga gagcgttgcg caactctagc tttaaacgcc cctgtcccc 8220
tcaacttgct tccccagcc cctctgattt acagattctg cagtccccga ggggtgcgcc 8280
tacgataccg aactcgcgg cagcctgcga ggcgagtatg atcgtcccat tttcggagt 8340
agcaactaa ggttcagaga ctactatgtc ccaggtcgg ctggtttgaa ggtccgctt 8400
cctctccctc cgccagcggg cgtgcgagg gactgggcga ggcagcgctt ccctaaggag 8460
gcgacccgca gccccggccc cctcccgact ccgccccgtt gcagggcccg ggtcggcgag 8520
gcctctcagc tctaagccc acgggacttg gtgattggg aggacggaag agctgggtg 8580
ggctttccac cagcggagaa agtctagtgg gcgtggtcgc gacgagggcg tggcctggtg 8640
ccccgcccc gtccgcgcgc tcaaagtga ggggtggtgt gggggcgggg tcagaacact 8700
ggcggccgat cccaacgagg ctccctggag ccgacgcag agcagcgccc tggccggggc 8760
aagcaggtat cgacaccgc gcggggcgctc ttgggctgga ccaggcgggc gcccggggc 8820
tgctgaggac cacaaaggc actgggggtc gtggtccagg ctgtgcttcc tcccgctgc 8880
cctggccct gcctccgccc ccgccccgc cttcctgccg ctaagccggc tgcggcgggg 8940
ccgattggcg cctgcggct tctgcgccg gggccagtct aatgcatggg gcccgggcg 9000
gggactaagg ggaaactgag tcacgtcgg gtgggagcag ttctgtgtg gaggcaccac 9060
ccccactgg gctcggggaa ggatccccct ccaagctatg cttgagggtc ccagcccca 9120
tctgtctcca caggggcgc accccactcc cgccttcccc ttcttcagca cccaggggtc 9180
ccgcccgtgc tccagcagc ctgcactggt ccggaatgg ctaggaggat ccgctgcagc 9240
cgctccctc cctccctc cctccctc cctccctc cctccctc cctccctc 9300
cccctcgcgt cccaagcccc cgtgtgctcc ctcgctggc tctccgcaca gtgtcagctt 9360
acacgcctta tatagtcga gcaggctcca gccgcggcct gctgccggga cctggggggc 9420
ggggagagga gagccggccc ctgactcacc cggaccgccc gaggtccag gctggcttgg 9480
ggggaggccg cgccagttta gtccctcggc ccaccttg ttgcaaagaa cctcaagcct 9540
ggattcaggc acccctcacc gttccagtc caaggggagg ggggctgctc ctgtcttcc 9600
aaagtgaggt ccgcccagca gcagcccagg ccagcctgac aaaatacctg cctcctatgg 9660
cttgggcgtg ctcaggggt gccggtgct gctggcccc tgtccaaggc tggatcctg 9720
agctggcccg gcctgcctgc ctgcccgc accatgctgg cactcacct tctctctct 9780
cctctcagga gccgcatca tggattcctt caaagtatg ctggaggggc cagcaccttg 9840
gggcttccg ctgcaaggg gcaaggactt caatgtgcc ctctccattt cccgggtgag 9900
cctaggtttg gggaggggg tccccagcg gtctttcgg gcttaggtct ccagaggtg 9960
atggggggag tctaacagg agctggtcag gggccagcag gccaggagat gtctaggctc 10020
ggagatgtag tggtaacctg ctgccacaag gactccaat gaggtggata ctgggaggga 10080
gcacccaggc ttctccagcc ctgcactgta cccgatgctg ttctcccaag ctctgtggc 10140
cacctctgag ggctggagg aggtcattg tgcaggatgg gagcctaaca tttcaggagg 10200
tatctaaact tgaggtggca atgcttggag ccaggccca ggcaggacac tgtgactata 10260
ggatttcact tcagcctcac tgccgcccag ggaatagcaa tcctcatccc gtttttccag 10320

```

```

atgagagaag aactcatgga gaggtggcgg ggctcgctca tcgagtcctat ggtgaagcag 10380
ggattggaat tgaggcacag catggcgtag attttttggt ggtagaaggg gtctctcccc 10440
agcctatgta aggacccaca tccactgttc ccattcagga tgtggtggcc tttagaccca 10500
agcagaagtg taggacaggg ctccattcta ggggcttaac ttcagcttcc aagagcctgc 10560
cctggtgtgg gtggagctgg aggctggctc ctccctgtag cagggggatt gccttataag 10620
cccaagaatg cagccccacg ctgggatggc caacagtggc tgcggtctgc agagctgaaa 10680
agggctggcc taggcctggc cccctgaacc ccactggtgg gcctctcagc tggtcaccag 10740
gctgcagctc cagctgtatg gtccagtgtg gagacacaac aaattgcctg cccagagtgg 10800
gtgaggccag cctgtcggct ggcatctctg actggcctgg gggtcaggag ggggtgggga 10860
cttctgccc ctatatccgc ctgcccagag agaccaccc aggcgcggg tgggcaggca 10920
gctgttgta ggaagcccaa ggcaagccca gcctggaggg gccagaggg tcgtggcctg 10980
aggaggggct caagctggag tctgtctgta ggagctgggc gtgggggtta ggggtggcag 11040
gccagcagtg ctcttctcag gggctccttg atggcattct cctggaacct gccccgccag 11100
cagggtagtg aggcagtggg tgcctatga cacacgtccc actacatagc cctcacacag 11160
ccctgaaacc tacctgacgt cctgtccctt gggaaagtgc tggcccagtg tgtctgggga 11220
gcctgaacct cagtttcttc cctgatggag atgactttca gatatggcct gttgggggca 11280
ctccgggctc cagctccctg gtcagcatcc ctggcatgtg ggcggggcca ctagctgac 11340
ccagccctgg agttggacct gggcccacat ggggtgggtga ggtgggcttt tctgagttag 11400
gccagccccc tccccctccc ctgacccag aatggagggg ggtgggaggg gcaagggtctg 11460
gctgtggggc cagggcctgg agatgaggta acgtctggga ctggggggct gggctgctca 11520
ggctgactca cccccacctc atgcagggtc cagccccctg gctttttccc tccttggttc 11580
ctctggcctt accctgcccc tggttgagc cctccctgc ctctctccag ccaccgccc 11640
agcgtgtct tctgtctctc tgctgcccc cccacgtctt gaacaccctt cactctctgt 11700
gcttctgcc ctctcactc tgggaaggga agcgtcccc gccccccacc cctctccag 11760
gagccagcta gctgcacccc aagaccccca cctcgggctc agcccacagc tcccaggagc 11820
cagccctgtg ggcagggagt ggctgggcca ggtttccctt ctactgactc accatgacct 11880
tgagttaagt acttccccctc tgggggtgtca ctccccata cacagtataa ggggttgatt 11940
tagttggatt gaactaaagg tgagggagtg gctcagggtg tctccaggtg ggctgacccc 12000
tcagtggggc ccccatgctc agcagaggtg gccacagtg gtggagcctt agggtcagag 12060
acacttcctg gctctgctc ttactagctg ggtgacttga ggcaagtgtt ttaacctctc 12120
tgtgtacatt tgcaagtga aaatgggtaa aatcccagat tactccacaa ggttggttga 12180
agattcagtg tcaatatgta gcatagttgg tgctcaataa actgaagcaa gtcttcttat 12240
ttagcgagtg aggaaggggc cgccagctc tcttagctt ctgacctcct acgcaagcaa 12300
gaggtcaagt tgagccagc tgcctttct tttccagtg ctgtcaagct ctgtgctgg 12360
ctgcctgcc ctctgacatc tctctgaaac ctcttgctc cctctcctt gcctcagctc 12420
agtctgtgca ctgaccacc tgaggagcct cctggggcca ctggcagcct ggaccccccc 12480
agatcccccc caccagtgta aattgtcttc cagcactgcc tcacaaaagc ctacttgatg 12540
cagtggcaag cctcttgcca gatggctggg tggctccctt ggcttggaac cagtcaagct 12600
gccctgctg tgttgctggg gctgggctag aggcctggaa ggggtttatc agggtcaccc 12660
tctcagggcc tgggagatac ccaatcccag acattaaaac tgccagtagc cctctacct 12720
tcaaagccaa gtccctggctc ctccccctgg cattcaaagc catcgtaagt gaactctcac 12780
ccgctaggca gcacacgcca ttctccttta ccgaggccca ccgcttctc aaagtcattc 12840
ctgatggtct cagctcatgc tgggtggcagc catttctccc agcctactgt ctctactcat 12900
tgccacagga accagggaact cccagctcaa gagcctgaag gattggggtc aggggaaatt 12960
ggcagtcgag ggcttgggag tgacagccat gtatggccta cgaagtccca gctgtcaact 13020
taggtcccat tcaggcagtg ttccagggga accgggagat aacagggcct gttcctggct 13080
ctcaaagggt cccagcagac ccctatagat ggcccccgag aggggtgctg ggggtgagag 13140
gtccataaga gcccccggtg gtttcgggga ggaagctgcc cctgcatgg gccagaggc 13200
atatctggta ggtggagtgg cctgggcagg aggccagcag gagcctcaaa aggcaatggt 13260
cctcctgaaa cacttgggct ttagcctgag cgtggctgtt tgtggacatc atagcaattt 13320
ctggactgtg ggggaggggt gtggcggtga atagataagc atcgtgactg gggaagctca 13380
ggtgagcacc acctgaggga gagggtctgg cagtgaataa ataagcagtg tgactgggaa 13440
attgtgaagc tcaggtgagc gccaccacct cctgggttgc ttagtgtcc agcagctgcc 13500
tagaactatg ttgaatgaag agctctctgg gttctggaag tgggacagct ttgggtgggg 13560
cagtggtacc accgtcagcc tggcttgggt ctgcagggtc cagggcctcg gtcactttgc 13620
ttctctctcc acagctcact cctgggggca aagcggcgca ggccggagtg gccgtgggtg 13680

```

actgggtgct gagcatcgat ggcgagaatg cgggtagcct cacacacatc gaagctcaga 13740  
 acaagatccg ggccctgcggg gagcgccctca gcctgggcct cagcaggatc gcgggtggac 13800  
 atggatgggt gcgcccgcgc tggcagtggg gatccctgcg gcccggcccg ctgtcacgct 13860  
 ttccttctcc tccagggccc agccggttca gagcaaaccc cagaaggtag gaggtgggc 13920  
 gggacatccg ggcggtgggc ggtgtgggct tggacggcca ggccctgctc ccctcctggc 13980  
 acattctcgg taccccaatc cctggccggg agtggagggc agaaaccgga gctaaggcgg 14040  
 gtctagggcc ctggagttga gccaggggct gctgcacggt cctggcacca cgcattgtccg 14100  
 cctgtctgtc cgccctgtctg tccgcctgct gcctcccgc gccggcgctg cgtgctcgcc 14160  
 cgcactcggg cagccctcgg tccctgctgg actgagatcg ccactcccaa atggggccct 14220  
 tgaacactga gtctgtctct ccccgtagcc tccaaataga ttaggggggt ggggtggggg 14280  
 tggggggctg gagctgccgc tgtcctctgc tgcaggcgcc ccacttccac ccaggccccc 14340  
 accttacctt gcccgcccgc cctgcccggc tgtgtctctg cccaggcctc cggccccgcc 14400  
 gcggaccctc cgcggtacac ctttgcaccc agcgtctccc tcaacaagac ggcccgccct 14460  
 ttgggcgccc ccgcccgcctg acagcgcccc gcagcagaat gggtagctcg gccctgccc 14520  
 gcccgcccgc accgcatcag gccactgtg gccccacgcc cgctgccgc tgetgctcag 14580  
 tctgtgctgc gcccagccc ggcggaaccg tgcggcacgc cccctggcg ccgggggtggg 14640  
 gctgcaggca cagggccctt cccgaggtg tggcgcttg cagggcaccg cctggggagg 14700  
 ggtctctgaa tgacgcgcg cccctgctg gcggtgggg gttgggtgt ggtgtcgggc 14760  
 cagctgagcc ccagacactc agtgcgcct tctccccgc tgttctgacc cctcccctc 14820  
 tttcttctc tctgtgtct tccctttgt ccttttatct gtctgtctgt cttatttct 14880  
 tcacaggtgc agaccctga caagtgcagt agccccctc tgccctgtgc tttcttctc 14940  
 cttttggcac tctgggtggc ggccccctc caccctggct gccctcctc ccacttcgc 15000  
 ctctgtctc ctacactacc cgccagcag ggtcctggc ctacacctta ccactccct 15060  
 cccatcactg taacccaaac ccacatgcac caaatcctg gaggggctgc cccaccgcc 15120  
 cacccccagt gtgggttct gagccacacc ctccccacag acagccgctc cgaccgctg 15180  
 tcccagatgc cagcaagcag cggctgatgg agaacacaga ggactggcg ccgcgcccg 15240  
 ggacaggcca gtgcggttcc ttccgcaccc ttgcccacct cacaggcacc gagttcagta 15300  
 agtgccagcc cagggcaggg ggtactttcc tgcccccag cccaggcgtg atccctgacc 15360  
 ctgtgtcttt tttggtcaat gcctgcctct gctctctcag tgcaagaccc ggatgaggag 15420  
 cactgaaga aatcaaggta cagggacggg caccagcccc tctccacct cctgcctct 15480  
 ccattccagc tactgcctg tgtctactcc tgaggtccc agctggggt ctcaattctc 15540  
 ccttccctc ttccttctt ccttccctc ttccttctt ccttccctc ttccttctt 15600  
 ccttccctc ttccttctt ccttccctc ttccttctt ccttccctc ttccttctt 15660  
 cctgcccccc ttccttctt ccttccctc ttccttctt ccttccctc ttccttctt 15720  
 gggaaaagta tgcctggag ctgcagagcc cagcctacac ccgctcccg gactggcacc 15780  
 accagcgctc tgcccagtg ctcaacgtgc agtgcgtagc cggccctctc cagccggctg 15840  
 cctctctgct cctccctctt ctgttccctc tgcccagggc acccccttag tgccctcagc 15900  
 ttctgcttac ctaccccc ctttctgtgc cctggcctga gcctcctgct ggccctggcc 15960  
 tggccgcca cctgggttca tctgacactg ccttccctct ttgcccctg gtactgctgt 16020  
 ctgccaggtc tgtgctgcct tgggcatgga ataaacattc tcagccctgc ttgctctgcc 16080  
 tgtcttctat ctttgtggac ctggtttgca ttgggggtgt ggggtgttt cgtggttcgg 16140  
 actgtttggg cctgcccgtc cttgttttca gtgggagggg gtacctggca aaggggccct 16200  
 gccctgccat cacagatggc ttcctggcat gaggggagcc ccaggagctg cctcagaagc 16260  
 gggagccctg cctcgtctcc cagctagaga ccgcacacca gctaactgga cattgctagg 16320  
 agaagctgcc cttcccctc ctaccccagt gggacctgga atccaactc gcagtttcca 16380  
 cgccccagc catctccgt ggggcccagc ggacccagg tgggggggtg ggccatgtca 16440  
 ggaagctcag ccatgcagg ccttgaatgg cagatcctgc agccagggtc ccaggacaga 16500  
 agccccagc ccagcctcat ctacacccca ggagccctg ggtggtgaga gggagtgggc 16560  
 tggggcctgg gcaagggtgg gcagcctcca ggggcatggg ggtggtgggc tctctcagc 16620  
 tgccctgggg tccacccccg tcccttgggg tccctgggca cccctttaga gtcacttctc 16680  
 ccggcaggcc ctaccgcccc cagccctacc agccgcccgc cctgggctgt ggaccctgcg 16740  
 tttgcccagc gctatgcccc ggacaaaacg agcacagtgc tgaccgggca cagccagccg 16800  
 gccacgcca cgcgctgca gagccgcacc tccattgtgc aggcagctgc cggaggggtg 16860  
 ccaggagggg gcagcaacaa cggcaagact cccgtgtgtc accagtcca caaggtcatc 16920  
 cgtgggtgg cctgttccct tccgacctg gcttcccat cctgcagccc agccccacct 16980  
 gtctgcccac ctgtcctgcc tcagctgcga ctggggggaa taaggattca gttctcagct 17040



```

ggagtaggag tagggacctg ggctgggtcc tccattctt aatcccacgc tacctacccc 17100
agcccaccca caacaactgc tagcagcatc tgccgtggcg aaatagccga agggccaacc 17160
ataggctgaa gctgcacccc tacctttgct gctctctggg caaagagggg cctgccccct 17220
cccagcgcgt ctgccccctc ctctgtctct ctgtctccct ctgctctcag agcatacagg 17280
cctggagcca ctccctctgt gcactgcccc gtggggccaa gcagcatcaa acacccccca 17340
gcacagcgt gccggattct agagccttcc taattcgag gcctggcctg ctctcatctc 17400
tgtcagctct tttttttttt tttttgaaac agagtctcac tgtgttggcc acgttggcgt 17460
gcagtggcgc gatctcggct cactgcaacc tctgcctcct gggttcaaga gattctcctg 17520
cctcagcctc ctgagtagct gggattacag gcacccgcca ccatgcctgg ctaattttgt 17580
atttttagta gagacggggt tttaccatgt tggccaggct ggtctcaaac tctcacctc 17640
aggtgatctc aggcctgcct tggcctccca aagtgtctggg actacagggtg tgagccactg 17700
tgcccagccg actctatcag ctcttgccag gtagaacagg caggccagca ggacaggcca 17760
gctccagggt ttgcccaggg gcggctcagc ttttatgagg ctccagtcgt cagcccttcc 17820
tcccggggtc ctccctgtc taaagctgcc tctcctgtca ccagcagttc agtgtggcgg 17880
actggctctg taagcttcat ggctgccacg gtcacttccc aagcctgtct tctatcctat 17940
gtgaaaaatg gggagaatga actgtccctc ccaaggcctc ctggtgggtg gtcagtcaac 18000
ctgaaggggg ccaagacccc cacctctctg cgtgtgtctc ctctgaccgc tctgcctcc 18060
ctgcaggggc cgctacctgg tggcgctggg ccacgcgtac caccggagg agtttgtgtg 18120
tagccagtgt ggaaggtcc tgggaagggtg tggcttcttt gaggagaagg gcgccatctt 18180
ctgcccacca tgctatgacg tgcgtatgc accagctgt gccaaagtga agaagaagat 18240
tacaggcgtg agtagggctg gctggcgggg aggtgggtccc aagcctgtca gtgggaacga 18300
gggctgctgg gaaaccacac gtccaggtct ctcccaggt gagcctccgg gtccttacca 18360
gcgtaataaa tgggctgctg tactggcctc accctgcatt agtcaggatg ctcttaacaa 18420
atgacatgt tctgtctcag aaaccgcca aggtgcaaa gagcaggagg accaagccag 18480
gagaagccct gggccctcct gactccact ttgggtctc cctgccctgg tgaaatgaca 18540
gaacggccaa cttgacacgc tgaagctgct ctgtctcatg cgtcctcctc atttctggat 18600
ccagagccag ggctgccagg agtagccaga gagctctgtg tggatgatgt catattagtg 18660
aggtttacct tgaccacgag cagtgggaaa ctcaaaataa tgggtggctta tttctcatct 18720
aaaaacatcc cggggtgggt ggtctgggac tgatctgggt gaccagggt ccgccttgtt 18780
gcttgactgt tggcagcacc tgcttactta ccactcatgg tgcaagatga cacttcagcc 18840
tccgcaaaaa gtctacctt ccagccagca ggaagtcgga aggagaagaa aggggacaga 18900
gccccatggc gtccatcctt agaggatgct gccacctgaa cctctgcttt catcctgttg 18960
gtcagaaccc agtcacatga ccacaccag tggcaacgga ggctgggaaa tatagtcttt 19020
attttgggca cccatgtgtc cagcaaaact ggggggtcca tcagtcggca agaacgggag 19080
agtgccgat gcagtggctg atgcttgat cccagcactt tgggagggtc aggtgggcag 19140
atcacctgag gtcaggagt caagaccagc ctggccaata tggtgaaacc ctgtctctac 19200
taaaaaataa aaaattagct ggggtgtgtg ggcacctgt agtcccagct acttgggagg 19260
ctgaggcagg agaatcgctt gatcttgaga ggtggagggt gcagttagcc aagattgtgc 19320
cactgccttc cagcctggga gacagcaaaa aaaaaaaaaa aaaaaaaa aaaaagggcc 19380
aggcacggtg gctcacacct gtaatcccag cactttggga ggccgagatg ggcgatcac 19440
gaggtcagga gattgagacc atcctggcta acacggtgaa accccatctc tactaaaaat 19500
acaaaaaaat tggccgggca tgggtggagta gtcccagcta ctccggaggc tgaggcagga 19560
gaatggcgtg aacctgggag gcagagcttg cagttagccg agatcgccg actgcactcc 19620
agcctgggca acagagcgag actctgtct caaaaagaaa aaaagaaaga gaaatctgcc 19680
tcccagcctt gggctcctgc cctaccagcc cacaccctg gtagagcctc ctctcccacc 19740
agctcaaagc ccaagttcct tcactgtgac cttgtctgct cctctaaaac aggcaacacc 19800
agacagttag aagagccagc cagacatggg cagaaaaacct attctgtga tctactggct 19860
gtgtgagcag gggctagtgt ctctctctgg gcctcactga agagaagggt ggacttatgc 19920
tagggccggc acggttgcaa ggtagatgta agatggggta cagggtgtgt ggagggcaga 19980
aatgcacat ccgaaggcta catgtcccc acacttatgt cttgcttggc ccacactgtt 20040
tcattttaaa atcagtagca aacaatttaa aaaatcagaa gatttgctg catgatgcag 20100
tggctcatgc ctgtaatccc agcactttgg gaggccaagg tgggaggatt gcttgagccc 20160
aggagttcaa gaccagcatg ggcacatag caagaccct gtttctacaa aaaaaaaa 20220
attagaaaat tagccaagt tgggtggcatg cacctgtggt cccagctact tgggaggcag 20280
agggaaagt agatctcctg ctttttatt ctttatgtat aatgatagg tcttgcctg 20340
ttgcccaggc tggagtgcag tggcatgatc actgctcact gcagccttga tctcctgggc 20400

```

tcagaggatc	ctcccacctc	agcctcccaa	atagctagga	ctagagggtgc	ccaccagcat	20460
gctcagcaga	tttttaaatc	tttttgtaga	gatgaggttt	tgctatgttg	cccaggctgg	20520
tctcgaactc	ctggcctcga	gcgatcctcc	caccttgggc	tcccaaagca	ctgggattac	20580
agacgtgagc	cactgcgccc	agcagatttc	tctttaacac	ctagatttca	gcctgagcca	20640
ggcaggcatt	cctgaatgaa	ccagtagtac	tgctcccaga	agaagaggtc	ctcctccgtg	20700
tgacacagtc	cccacttggc	ccttgccagg	attggatctg	ggatccctgg	atttaaactc	20760
agggccatcc	tcataacagc	ctcacaaggc	tgggattagc	ttcccagttc	acaagggaag	20820
aaaccaagac	ttgagaaggt	caaggctctg	ccagaccac	acatcttgga	ccctcatacc	20880
gcctcgaggc	cccattgctg	cctctgcctg	ctccagatgt	gaatactgct	ggccctggct	20940
ggccccggct	ggccccgagg	gtcctaggga	tgaacagccc	agcccaggga	gagctcagcc	21000
ccttgtgcct	ctgccccctc	ccacctcctg	cggaggccag	tcgactcacc	cacaaagggc	21060
caggcactgt	ggggatagat	cagctaacaa	aacagttgat	gcttcctgcc	cttctggggc	21120
ttacattttg	gctggaagaa	gaggggagag	gcagactgta	agcaataagc	gcaataagta	21180
ggttgccctg	aagtaatgtt	agatcacgtt	acggaaaaca	ggaaagagca	gagcgacaag	21240
tgctgggggtg	cgtggtgcag	ggaaggcagc	tggtgctgct	tggtgtggtc	agagtggggc	21300
ctcatggaga	agactgcatt	cgagcagaaa	cttgaagggg	gtgaggggtg	agcctagaga	21360
tatctggggc	agagcagtc	aggcagaggg	gacagccggt	gtcaagcccc	ggacaggagt	21420
gtgcctgggtg	tgccagtctc	aggcaagagg	ccagtgtgca	gaggcaaggt	gagaacgcaa	21480
gggagagcag	tggcgagac	gggtgggaac	cgagtcagac	ctgctggcct	ccagcctctg	21540
catgggggctt	ggctcttgct	gggagcaatg	ggaagcagta	cacagtttca	tgcaggggga	21600
gaaggcctgt	cttgggttgc	aggggcacgc	tgtggcagct	gggatcagag	agaggagctt	21660
gtaggccagt	tggtatgtgg	tcccacgggc	cagatggcca	tggcttacct	cacttcaggg	21720
aggctgtgag	aagcactcag	aatctggatg	tgcttggggg	gtgggcccc	ctggatttcc	21780
tggtggacct	ggtgtggggg	gtgagaggag	ggtgtgtttg	gctgcagcag	acaggagaat	21840
ggagttgcca	tccgcgtgat	ggggatggct	gtgggaggag	aggtttgggg	tgaggggaatc	21900
aggaactgag	tgctggacat	ggcaagtctg	aaggcgagct	ggtcgtccac	tcagagacct	21960
tggagttgga	gatggagggtg	tgggagtcct	gaacagttag	atgtagtggt	taccgcgaga	22020
aggaacaggg	cttgcggcca	gccctcctgt	gttcccgtga	cccagggcag	ggcaggaggg	22080
gcctgagcct	gccgagtgc	tgggacctcc	ttccaggaga	tcatgcacgc	cctgaagatg	22140
acctggcacg	tgcactgctt	tacctgtgct	gcctgcaaga	cgcccatccg	gaacaggggc	22200
ttctacatgg	aggagggcgt	gccctattgc	gagcgaggta	cccactggcc	agtgagggtg	22260
aggagggatg	gtgcattggg	caggcatgaa	tccaggtcct	ctttctctct	gccccattc	22320
tcagactatg	agaagatgtt	tggcacgaaa	tggcatggct	gtgacttcaa	gatcgacgct	22380
ggggaccgct	tcctggaggc	cctgggcttc	agctggcatg	acacctgctt	cgtctgtgcg	22440
gtgagagccc	cgccctcga	actgagcccc	aagcccaccg	gccctctggt	cattccccag	22500
gagatgcagg	agaagtggg	aaggggcctc	tcctgctgcc	cccaacccca	tgtgactggg	22560
cctttgctgt	ccttagatat	gtcagatcaa	cctggaagga	aagaccttct	actccaagaa	22620
ggacaggcct	ctctgcaaga	gccatgcctt	ctctcatgtg	tgagccccct	ctgcccacag	22680
ctgccgcggt	ggccccctag	ctgagggggc	tggagtcgtg	gccctgcatt	tctgggtagg	22740
gctggcaatg	gttgccctaa	ccctggctcc	tggcccagac	ctggggctcc	ctggggccctg	22800
ccccacccac	cttatcctcc	caccccactc	cctccaccac	cacagcacac	cgatgctggc	22860
cacaccagcc	ccctttcacc	tccagtgcga	caataaacct	gtacccagct	gtgtcttgtg	22920
tgcccttccc	ctgtgcatcc	ggaggggcag	aatttgaggc	acgtggcagg	gtggagagta	22980
agatggtttt	cttgggctgg	ccatctgggt	ggtcctcgtg	atgcagacat	ggcgggctca	23040
tggttagtggt	aggaggtaca	ggcgagacct	catgtgccag	gcccggtgcc	cacagacatg	23100
aggggagcca	ctggtctggc	ctggcttgga	ggttagagaa	gggtagttag	gaagggtagt	23160
tagcatgggtg	gctcatgcct	gtgatcccag	cactttggaa	ggccaagggtg	ggcagatcgc	23220
ttgaggtcag	gagttcgaga	cctcatggcc	aacacgggtga	aacagcgtct	ctagtaaaaa	23280
tacaaaaatt	agccgagtg	ggtggggcat	gcctgtaatc	ccagccactc	aggaggctga	23340
ggcgggaaaa	tcacttgaa	ctgggaagtg	gaggttgtag	tgagctgaga	tcacaccact	23400
gcgcgcgagc	ctgggtggca	gatggcagag	cgagaccctg	cttcaaaaaa	aaaaaaaaaa	23460
aaaaaaaaaa	gaagggtagt	tgtagtggg	ggtggatctg	cagagatatg	gtgtggaaaa	23520
cagcaatggc	cacagcaaag	tcctggaggg	gccagctgcc	gtccaaacag	aagaaggcag	23580
ggctggagag	ggtagccctt	aggtcctggg	aagccacgag	tgccaggcag	tagagctggg	23640
gctgtctctt	gaggttaggg	cagggcaagg	cacagcagag	tttgaatatg	gtttgtgttg	23700
tattgcagaa	aagaggcccc	agaacactga	gggagtgtag	gagggagggt	gggaggagga	23760

```

gttgccagcag ggcctagggg cggggggccag gcaagggagg ggcagagagt aatatggcag 23820
agatgggacc cagtggcagg tccgggggat gagggatgga gagaaggaca ggagcgttgc 23880
caggcatctg gcctatacca gacatgctca cgctgtctcc cgcgaaacct ctagcaacct 23940
tgcgccgttg tctgcaatca cttatttcat tttttctttt ttaactttaa tttttttgt 24000
ttttaagaga caggatctcc ctaggttgcc cgggctgggt tcaaactcct gggctcaagc 24060
aattcttcct ccttagcccc aaagtgctgg cattacagg gtgagccacc atgcctggcc 24120
cacttatttt ctagatgagg cacagaaaga ttgggagact tgaccaaggt cacgctgtca 24180
ttgagccatg agccagacta gaatccaggc ctgaagctgg gtgcgctgtc ccaggactgg 24240
ctggcactga gtaccatttg ccagcgagca tctctctggg aagctgactt ctgcccggta 24300
cctggaggac tgtagacctt ggtggtggcg ccgtcactct ggggcttcct gcctccact 24360
gatgcccgca ccaccctaga gggactgtca tctctcctgt cccaagcctg gactggaaag 24420
actgaagaga agccttaagt aggccaggac agctcagtgt gccatggctg cccgtccttc 24480
agtgggccct ggcattgagga cctgcaacac atctgttagt cttctcaaca ggcccttggc 24540
ccggtccctt ttaagagacg agaagggtg ggcacggtga ctcacacctc taatcccagc 24600
actttggaag gctgaggctg gagaagggt ccagcttagg agttcaggac cagcctgggc 24660
aacatggtga gaccctgttt tgttttgttt tttgtttttt tgagatggag tcttgcctctg 24720
tcgcccaggc tggagtgcag 24740

```

<210> 42

<211> 25

<212> DNA

<213> *Rattus norvegicus*

<400> 42

gcactacctt gaaggaatcc atggt

25

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/11664

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : 435/320.1, 325; 530/350, 388.1; 536/23.5; 514/44

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 325; 530/350, 388.1; 536/23.5; 514/44

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BODEN et al. LMP-1, a LIM-domain protein, mediates BMP-6 effects on bone formation. Endocrinology. 1998. Vol. 139, No. 12, pages 5125-5134, entire document.	1-37
A	HOGAN et al. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes and Development. 1996. Vol. 10, pages 1580-1594, entire document.	1-37
A	KINGSLEY, DM What do BMPs do in mammals? Clues from the mouse short-ear mutation. Trends Genet. January 1994. Vol. 10, No. 1, pages 16-21, entire document.	1-37

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

16 JULY 2000

Date of mailing of the international search report

29 AUG 2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

ANNE-MARIE BAKER

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)\*

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/11664

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

A61K 48/00; C07H 21/04; C07K 14/00, 16/00; C12N 15/63, 15/85, 15/86

## B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST

Dialog (file: medicine)

search terms: gene(w)transfer, LMP, bone, LIM, mineral?, human, osteogen?

## SEQUENCE LISTING

<110> Bodan M.D., Scott D  
Hair Ph.D., Gregory A

<120> LIM Mineralization Protein Splice Variants

<130> splice variant LMP

<140> 6148.0223-00304

<141> 2000-04-29

<150> 60/132,021

<151> 1999-04-30

<160> 42

<170> PatentIn Ver. 2.1

<210> 1

<211> 457

<212> PRT

<213> Rattus norvegicus

<400> 1

Met	Asp	Ser	Phe	Lys	Val	Val	Leu	Glu	Gly	Pro	Ala	Pro	Trp	Gly	Phe
1				5					10					15	

Arg	Leu	Gln	Gly	Gly	Lys	Asp	Phe	Asn	Val	Pro	Leu	Ser	Ile	Ser	Arg
		20						25					30		

Leu	Thr	Pro	Gly	Gly	Lys	Ala	Ala	Gln	Ala	Gly	Val	Ala	Val	Gly	Asp
		35				40						45			

Trp	Val	Leu	Ser	Ile	Asp	Gly	Glu	Asn	Ala	Gly	Ser	Leu	Thr	His	Ile
	50					55					60				

Glu	Ala	Gln	Asn	Lys	Ile	Arg	Ala	Cys	Gly	Glu	Arg	Leu	Ser	Leu	Gly
65				70					75					80	

Leu	Ser	Arg	Ala	Gln	Pro	Ala	Gln	Ser	Lys	Pro	Gln	Lys	Ala	Leu	Thr
			85					90					95		

Pro	Pro	Ala	Asp	Pro	Pro	Arg	Tyr	Thr	Phe	Ala	Pro	Ser	Ala	Ser	Leu
			100					105					110		

Asn	Lys	Thr	Ala	Arg	Pro	Phe	Gly	Ala	Pro	Pro	Pro	Thr	Asp	Ser	Ala
		115					120					125			

Leu	Ser	Gln	Asn	Gly	Gln	Leu	Leu	Arg	Gln	Leu	Val	Pro	Asp	Ala	Ser
	130				135						140				

Lys	Gln	Arg	Leu	Met	Glu	Asn	Thr	Glu	Asp	Trp	Arg	Pro	Arg	Pro	Gly
145					150					155					160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175  
 Glu Phe Met Gln Asp Pro Asp Glu Glu Phe Met Lys Lys Ser Ser Gln  
 180 185 190  
 Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Thr Ile Pro Gln Glu  
 195 200 205  
 Ser Trp Pro Gly Pro Thr Thr Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220  
 Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240  
 Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255  
 Asn Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Gly Thr Gly Gly  
 260 265 270  
 Gly Ser Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Lys Ile  
 275 280 285  
 Ile Arg Gly Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu  
 290 295 300  
 Glu Phe Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe  
 305 310 315 320  
 Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro Ser Cys Tyr Asp Val Arg  
 325 330 335  
 Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile  
 340 345 350  
 Met His Ala Leu Lys Met Thr Trp His Val Pro Cys Phe Thr Cys Ala  
 355 360 365  
 Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly  
 370 375 380  
 Ala Pro Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys  
 385 390 395 400  
 Arg Gly Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala  
 405 410 415  
 Leu Gly Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln  
 420 425 430  
 Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Lys Pro Leu  
 435 440 445  
 Cys Lys Ser His Ala Phe Ser His Val  
 450 455

<210> 3  
 <211> 1696  
 <212> DNA  
 <213> Rattus norvegicus

<400> 2  
 gcacgagggat cccagcgccg ctcctggagg ccgcccaggca gccgcccagg cgggcattca 60  
 ggagcaggta ccattggattc cttcaaggta gtgctggagg gacctgcccc ttggggcttc 120  
 ogtctgcaag ggggcaaggc cttcaacgtg cccctctcca tetctegget cactcctgga 180  
 ggcaaggccg caccaggccg tgtggccgtg ggagactggg tactgagtat cgacggtgag 240  
 aacgcgggaa gcttcacaca cattgaagcc cagaacaaga tccgtgcctg tgggggagcg 300  
 ctacgcttgg gtcttagcag agcccagcct gctcagagca aaccacagaa ggccctgacc 360  
 cctccgcccg accccccgag gtacactttt gcaccaagcg cctccctcaa caagacggcc 420  
 eggcccttgg gggcaccccc acctactgac agcgcctgtt cgcagaatgg acagctgctc 480  
 agacagctgg tccctgatgc cagcaagcag cggctgatgg agaatactga agactggcgc 540  
 ccgcgccag ggacaggcca gtcccgttcc ttccgcatcc ttgctcacct caccggcaca 600  
 gaggttcctgc aagaccggga tgagggaattc atgaagaagt caagccagggt gccacggaca 660  
 gaagccccag ccccagcctc aaccataccc caggaatcct ggccctggccc caccaccccc 720  
 agcccaccca gccgcccacc ctgggcccga gatcctgcat ttgctgagcg ctatgcccc 780  
 gacaaaacca gcacagtgtt gaccgcacac agccagccag ccacacctac gctcctgacg 840  
 aacgycacct ccatagttca ggctgcagct ggagggggca caggaggagg cagcaacaat 900  
 ggaagacgct ctgtacgcca ccagtgcac aagatcatcc gccgcccata cctggtagca 960  
 ctgggccaag cgtaccatcc tgagggaattt gtgtgcagcc agtggtggga ggtcctggaa 1020  
 gagggtggct tcttogagga gaaggagct atcttttgc cctcctgcta tgatgtgcgc 1080  
 tatgcacca gctgtgcaca atgcaagaag aagatcactg gagagatcat gcatgcgctg 1140  
 aagatgacct ggcattgtcc ctgcttccc tgtgcagcct gcaaaaaccc tatccgcaac 1200  
 agggctttct acatggagga ggggctccc tactgcyagc gagattacga gaagatgttt 1260  
 ggacacaaagt gtgcgggctg tgacttcaag atcgatgccc gggacccgtt cctgggaagc 1320  
 ctgggtttca gctggcatga taagtgtttt gtttgcgcaa tatgtcaaat caacttggaa 1380  
 ggaagacact ctactccaa gaaggacaa cccctgtgca agagccatgc cttttccac 1440  
 gtatgagcac ctctccacac tactgcacc ctactctgoc agaagggtga taaaatgaga 1500  
 gagctctctc tccctcgacc tttctgggtg gggctggcag ccattgtct agccttggct 1560  
 cctggccaga tctggggct cctctccac agtcccttt cccacacttc ctccacccc 1620  
 accacogtca ctccacagtg ctagcctcct agcccagtt cactctggtg tcacaataaa 1680  
 cctgtatgta gctgtg 1696

<210> 3  
 <211> 260  
 <212> DNA  
 <213> Rattus norvegicus

<400> 3  
 ttctacatgg aggagggggc tccctaactgc gagcgagatt acgagaagat gtttggcaca 60  
 aagtgtogcg gctgtgactt caagatcgat gccggggacc gtttctcga agccctgggt 120  
 ttacagctggc atgatacgtg ttttgtttgc gcaatagtgc aaatcaactt ggaaggaaag 180  
 acctttctact ccaagaagga caagccctgt tgcaagagcc atgccttttc ccacgtatga 240  
 gcaactcttc acactactgc 260

<210> 4  
 <211> 16  
 <212> DNA  
 <213> MNLV



<400> 4  
aaagctttttt tttttg

16

<210> 5  
<211> 13  
<212> DNA  
<213> MMLV

<400> 5  
aaagcttgggt atg

13

<210> 6  
<211> 223  
<212> DNA  
<213> Homo sapiens

<400> 6  
atccttgcctc acctcaaggc caccagagtc atgcaggacc cggatgagga gcaactgaag 60  
aaatcaagcc aggtgcacag gacagaagcc ccagcccccag cctcatctac accccaggag 120  
ccctggcctg gccctaccgc cccagaccct accagccgcc cgccttgggc tgtggacct 180  
gggtttgcgc agcgtatgc cccagacaaa accagcacag tgc 223

<210> 7  
<211> 717  
<212> DNA  
<213> Homo sapiens

<400> 7  
atggattctc tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcaaggg 60  
ggcaaggact tcaatgtgcc cctctccatt tccgggtcct ctcctggggg caaagcggcg 120  
caggccggag tggccgtggg tgaactgggtg ctgagcatcg atggcgagaa tgcgggtagc 180  
ctcaaccaca tcgaagctca gacaaagalc cgggctctgc ggtagcgcct cagcctgggc 240  
ctcagcaggg ccagcctggc tcagagcaaa ccgcagcaag cctccgcccc cgcgcgggac 300  
cctccggcgt acacctttgc acccagcgtc tccctcaaca agacggcccg gcccttggg 360  
cgcccccgcc ccgtgacag cgcgccgcaa cagaatggac agccgtctcc accgctggtc 420  
ccagatgcct gcaagcagcg gctgctggag aacacagagg actggcgccc gcggccgggg 480  
acaggccagt cgccttctt cgcctcctt gccacctca caggcaccca gttcattgca 540  
gacccggatg aggagcactt gaaagaaatca agccaggtgc ccaggacaga agcccccagg 600  
ccagcctcat ctacacccca ggagccctgg cctggcctta ccgccccag ccctaccagc 660  
cgcccccgct gggctgtgga cctgngtlt gccgagcgt atgccccgga caaaacg 717

<210> 8  
<211> 1488  
<212> DNA  
<213> Homo sapiens

<400> 8  
atcgatggcg aagatgcggg tagcctcaca cactcgaag ctccagacaa gatccgggccc 60  
tgccggggagc gctcagcctt gggcctcagc agggcccccag cggttcagag caaacccgag 120  
aaggcctccg ccccgccgc ggacccctcg cggtaacact ttgaccccag cgtctccctc 180  
aacaagaagg cccggccctt tggggcgccc ccgcccgtg acagcgccc gcaacagaat 240  
ggacagccgc tccgaccgct ggtcccagat gccagcagc agcggctgat ggagaacaca 300

```

gaggactggc ggcgcggccc ggggacaggc cagtgcggtt ccttcggcat ccttgcccac 360
ctcacaggca ccgagttcat gcaagaccag gatgaggagc acctgaagaa atcaagccag 420
gtgcccagga cagaagcccc agccccagcc tcactctacac ccagaggacc ctggcctggc 480
cctaccgccc ccagccctac cagccgccc ccttgagctg tggacccctg gtttgccag 540
cgctatgcgc cggacsaacc gagcacagtg ctgacccggc acagccagcc ggcacccccc 600
acgcgcgtgc agagccgcac ctccattgtg caggcagctg ccggaggggt gccaggaggg 660
ggcagcaaca acggcagagc tcccggtgtg caccagtgc acaaggctat ccggggccgc 720
tacctggtgg cgttgggcca cgggtaccac ccggaggagt ttgtgtgtag ccagtgtggg 780
aaggctcctg aagagggctg cttcttttgg gagaagggcg ccactctctg cccacccatg 840
tatgacgtgc gctatgcacc cagctgtgac aagtgcacga agaagattac aggcagagac 900
atgcacgccc tgaagatgac ctggcaccgt cactgcttta cctgtgctgc ctgcaagacg 960
cccatccgga acagggcctt ctacatggag gagggcgtgc cctattgcca gcgagactat 1020
gagaagatgt ttggcaccga atgcccagtc tgtgacttca agatcgacgc tggggaccgc 1080
ttcctggagg cctgggcctt cagctggcat gacacctgct tctgtgtgct gctatgtcag 1140
atcaacctgg aaggaagac cttctactcc aagaaggaca ggccctctct caagagccat 1200
gccttctctc abgtgtgagc ccttctgccc caccagctgc ggggtggccc ctgacctgag 1260
gggctggag tctggccct gcatctctgg gttagggctg ccatggctgc ctaaacctg 1320
gctcctggc cagccctggg ctcccgggc cctgcccacc cacttatctc tccccccca 1380
ctccctccac caccacagca caccggtgct ggcacacaca gccccccttc acctccagt 1440
ccacatana cctgtacca gctgaattcc aaaaaatca aaaaaaaa 1488

```

&lt;210&gt; S

&lt;211&gt; 1644

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

```

atggattcct tcaaggtagt gctggagggg ccagcacctt ggggcttccg gctgcagggg 60
ggcaaggact tcaatgtgcc cctctccatt tccgggtca ctccggggg caaagcgggg 120
caggccggag tggcctgtgg tgaactgggt ctgagcatcg atggcgagaa tgcgggttagc 180
ctcacacaca tcgaagctca gacaaagatc cgggcttggg gggagcgcc cagcctgggc 240
ctcagcaggg cccagccggt tcagagcaaa ccgcagaag cctccgccc cgcgcgggac 300
cctccgpggt acacctttgc acccagcgtc tccctcaaca agacggcccg gccctttggg 360
ggcgcgcgcg ccgctgacag cgcgcgcgar cagaatggac agccgctccg accgctgggc 420
ccagatgcca gcaagcagcg gctgatggag aacacagagg actggcgccc gggcgcgggg 480
acaggccagt cggcttccct ccgcctccct gccacccca caggcaccga gttcatgcaa 540
gacccggatg aggagcact gaagaaatca agccaggtgc ccaggacaga agccccagcc 600
cagcctcat ctacacccc gtagccctgg cctggcccta ccgccccag cctacccagc 660
cgcccgccct gggctgtgga cctgctgtt gccgagcgt atgcccggg caaaacggag 720
acagtgtgga cccggcacag ccagcgggac ccgcccagcc cgtgcagag ccgcacctc 780
attgtgcagg cagctgcggg aggggtgcca ggaaggggca gcaaccacg caaagactcc 840
gtgtgtcacc agtgccacaa ggtatcccg ggcgctacc tgggtgggtt gggccacggg 900
taccaccccg aggagtttgt ggttagccag tgtgggaagg tctggaaga ggggtgcttc 960
tttgaggaga agggcgccat cttctgccc ccatgctatg acgtgcgcta tgcacccagc 1020
tgtccaaat gcaagaagaa gattacagga gatcatgc acccctgaa gatgacctg 1080
cagctgcact gctttacctg tgetgctg aagacgccc tccggaacag ggccttctac 1140
atggaggagg gcgtgcccct ttgcagagca gactatgaga agatgtttgg caggaattgc 1200
catggctgtg acttcaagat cgaagctggg gacgcttcc tggaggccct gggcttcagc 1260
tggcatgaca cctgcttcgt ctgtgagata tgtcagatca acctggaagg aaagaccttc 1320
tactccaaga aggacaggcc tctctgcaag agcctatgct tctctcatgt gtgagccct 1380
tctgcccaca gctgcgcggg tggcccttag cctgaggggc ctggagctgt ggcctgtcat 1440
ttctgggtag gctggcaat ggttgcctta accctggctc ctggcccgag cctgggtctc 1500
cgggcctctg cccacccacc ttatccctcc accccactcc ctccaccacc acagracacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgcacc aataaacctg taccagctg 1620
aattccaaaa aatccaaaa aaaa 1644

```

&lt;210&gt; 10

&lt;211&gt; 457

&lt;212&gt; PRT

&lt;213&gt; Homo sapi ns

&lt;400&gt; 10

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe  
 1 5 10 15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg  
 20 25 30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp  
 35 40 45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile  
 50 55 60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly  
 65 70 75 80

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala  
 85 90 95

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu  
 100 105 110

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala  
 115 120 125

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser  
 130 135 140

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly  
 145 150 155 160

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr  
 165 170 175

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln  
 180 185 190

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu  
 195 200 205

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp  
 210 215 220

Ala Val Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser  
 225 230 235 240

Thr Val Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln  
 245 250 255

Ser Arg Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly

260	265	270
Gly Ser Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Lys Val		
275	280	285
Ile Arg Gly Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu		
290	295	300
Glu Phe Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe		
305	310	315
Phe Glu Glu Lys Gly Ala Ile Phe Cys Pro Pro Cys Tyr Asp Val Arg		
325	330	335
Tyr Ala Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile		
340	345	350
Met His Ala Leu Lys Met Thr Trp His Val His Cys Phe Thr Cys Ala		
355	360	365
Ala Cys Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly		
370	375	380
Val Pro Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys		
385	390	395
His Gly Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala		
405	410	415
Leu Gly Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln		
420	425	430
Ile Asn Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Arg Pro Leu		
435	440	445
Cys Lys Ser His Ala Phe Ser His Val		
450	455	

&lt;210&gt; 11

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 11

gccagggttt tccagtcac ga

22

&lt;210&gt; 12

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 12

gccagggttt tccagtcac ga

22

<210> 13  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 13  
tcttagcaga gccagcctg ct

22

<210> 14  
<211> 22  
<212> DNA  
<213> Homo sapiens

<400> 14  
gcctgaactc tgtgccctg ag

22

<210> 15  
<211> 20  
<212> DNA  
<213> Rattus norvegicus

<400> 15  
atccttctc acctcaggg

20

<210> 16  
<211> 22  
<212> DNA  
<213> Rattus norvegicus

<400> 16  
gcactgtgct ggttttgtct gg

22

<210> 17  
<211> 23  
<212> DNA  
<213> Homo sapiens

<400> 17  
catggattcc ttcaaggtag tgc

23

<210> 18  
<211> 20  
<212> DNA  
<213> Homo sapiens

<400> 18  
gttttctctg gggcagagcg

20

<210> 19  
<211> 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: adaptor for  
Marathon RACE reactions

&lt;400&gt; 19

ctaatacagac tcaactatagg gctcgagcgg ccgcccgggc aggt

40

&lt;210&gt; 20

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: PCR primer  
specific for Marathon RACE adaptor

&lt;400&gt; 20

ccatcctaatacgaactcaact atagggc

27

&lt;210&gt; 21

&lt;211&gt; 765

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

```

ccgttggttg taaaacgacg cagagcagcg cctcgcccg gccacgcagg agccggcacc 60
atggattcct tcaaggtagt gctggagggg ccagcaacct ggggcttcgg gctgcaaggg 120
ggcaaggact tcaatgtgac ctctccatt tcccggtca cctctgggg caaggccgtg 180
caggccggag tggcogtaag tgactgggtg ctgagcatcg atggcgagaa tgcgggtagc 240
ctcaacacac tcaagctca gaacaagatc cgggctctcg gggagcgct cagcctgggc 300
ctcaacaggg ccagccgggt tcagaaacaa ccgcaaaagg cctccggccc cgcgcgggac 360
cctccggggt acacctttgc accaagcgtc tccctcaaca agaogggccg gcccttgggg 420
ggccccccgc ccgtgacag cgcgcccgag cagaatggac agccgctccg accgctggtc 480
ccagatgcca gcaagcagcg gctgatggag aacacagagg actggcgggc ggggccccgg 540
acaggccagt gccgttcctt tgcatacctt gctcacctta caggcaccga gtccatgcaa 600
gaaccgggat agggagcacc gaagaaatca agccaggtgc ccaggacaga agccccagcc 660
ccagcctcat ctacacccca ggaagcctgg cctgggctta cggccccag cctaccagc 720
cgcccgccct gggctgtgga cctgcgttt gcagagcgct atgac 765

```

&lt;210&gt; 22

&lt;211&gt; 1689

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

```

cgacgcagg cagcgccctg gccgggcca gccaggcgg gcatcatgga ttccctcaag 60
gtagtgtctg agggggccag accttgggg ttccggtgc aggggggcaa ggacttcaat 120
gtgcacctct ccatttccc gctcactcct gggggcaaa cggcgaggc cggagtggcc 180
gtgggtgact ggggtgtgag catcgatggc gagaatgcgg gtacctcac acacatcgaa 240
gtcagaaca agatccgggc ctgcggggag cgctcagcc tgggcctcag caggggccag 300

```

```

ocgggttcaga gcaaacccga gaaggeotcc gcccccgcgg oggaacotcc ggggtacacc 360
tttgacccca ggcctctccct caacaagacg gcccgggccct ttggggcgcc ccggcccgct 420
gacagcgccc cgcacacaga tggacagccg ctccgaccgc tggccccaga tggcagcaag 480
cagcgggtga tggagaaac agaggactgg cggccgcggc cggggacagg ccagtgcgt 540
tccttcgcga tccttgccca cctcacaggc accgagttca tgcaagacc ccgatgaggag 600
cacctgaaga aatcaagcca ggtgcccagg acagaaagcc cggcccagc ctcctctaca 660
ccccaggagc cctggcctgg ccctaccgcc ccagcccta ccagccgcc gccctgggct 720
gtggaccctg cgtttgcga ggcctatgcc ccggacaaa cagcacagc gctgacccgg 780
cacagccagc eggcacagcc cagcccgctg cagagccgca cctccatgt gcaggcagct 840
gcgggagggg tggcaggagg gggcagcaac aacggcaaga ctcccgctg taccagtg 900
cacaaggctc tcgggggccc ctacctggtg gcggtggg ccgcgtacca ccggaggag 960
tttgtgtgta gccagtgtgg gaaggtctct gaagagggtg gcttctttga ggagaaggc 1020
gccatctctc gccaccatg ctatgaagtg cgtatgcac ccagctgtgc caagtgcag 1080
aagaagatta caggcagat catgcacgcc ctgaagatga cctggcacgt gcactgctt 1140
acctgtgctg cctgcaagac gccatccgg aacaggccct tctacatgga ggaggcgctg 1200
ccctattgcg agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagategacg ctggggaccg cttcctggag gccctgggct tccgtggca tgacacctgc 1320
ttcgtctgtg cgtatgtcga gatcaacctg gaaggaaaga ccttctact caagaaggac 1380
aggcctctct gcaagagcca tgcttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcctgga gtcgtggccc tgcattctg ggtaggctg 1500
gcaatggttg ccttaacct ggctcctggc ccgagcctgg gctcccggg ccctgcccac 1560
ccaccttate ctccaccccc actccctcca ccaccacag acaccggtg tggccacacc 1620
agcccccttt cacctccagt gccacaataa acctgtacc cgtggaattt caaaaatcc 1680
aaaaaaaaa

```

<210> 33  
 <211> 22  
 <212> DNA  
 <213> Homo sapiens

<400> 23  
 gcactgtgct cgttttgctc gg 22

<210> 34  
 <211> 21  
 <212> DNA  
 <213> Homo sapiens

<400> 24  
 tccttgctca cctcagggc a 21

<210> 25  
 <211> 30  
 <212> DNA  
 <213> Homo sapiens

<400> 25  
 tcttcacccg ggtcttgcat gaactcgggtg 30

<210> 26  
 <211> 28  
 <212> DNA

<213> Homo sapiens

<400> 26

gccccggccc gctgacagcg ccccgcaa

25

<210> 27

<211> 24

<212> DNA

<213> Homo sapiens

<400> 27

tccttgcctc cctcacgggc accg

24

<210> 28

<211> 32

<212> DNA

<213> Homo sapiens

<400> 28

gtatatacgc tcaactatagg gc

22

<210> 29

<211> 33

<212> DNA

<213> Rattus norvegicus

<400> 29

gggctgatg gagaatactg aag

23

<210> 30

<211> 23

<212> DNA

<213> Rattus norvegicus

<400> 30

atcttgtggc actggtggca tac

23

<210> 31

<211> 22

<212> DNA

<213> Rattus norvegicus

<400> 31

tgtgtcgggt cagcactgtg ct

22

<210> 32

<211> 1620

<212> DNA

<213> Homo sapiens



&lt;400&gt; 32

```

atggattccct tcaaggtagt gctggagggg cccagaccctt ggggcttccg gctgcaagg 60
ggcaagggaat tcaatgtgcc cctctccatt tcccgggtca ctactggggg caaagcggcg 120
caggccgggag tggccgtggg tgaactgggtg ctgagcatcg atggcgagaa tgcgggttagc 180
ctcacacaca tcgaagctca gaacaagatc cgggctgtcg gggagcgccct cagcctgggc 240
ctcagcaggg cccagccggg tcaagagcaaa ccgcagaaag cctccgcccc cgcgcgggac 300
cctccgcggt acacctttgc acccagcgtc tccctcaaca agacggcccc gccctttggg 360
ggccccccgc ccgctgacag cgcgccgcaa cagatggac agccgctccg accgctggtc 420
ccagatgcca gcacgcagcg gctgatggag aacacagagg actggcgggc gcggccgggg 480
acaggccagt cgggttccct ccgcctcctt gcccacctca caggcaacga gttcatgcaa 540
gacccggatg aggagcaact gaagaaatca agccagggtg ccaggacaga agccccagcc 600
ccagcctcat ctacaccccc ggggccttgg cctggcccta ccgccccccg cctaccagc 660
cgccccccct gagctgttga cctgctgttt gccgagcgct atgccccgga caaaccgagc 720
acagtgtctg cccggcacag ccagccggcc acgcccacgc cgtgacagag ccgacacctc 780
attgtgcagg cagctgcagg aggggtgcca ggggggggca gcaacaacgg caagactccc 840
gtgtgtcacc agtgccscas ggtcatccgg ggcgcctacc tggtaggggtt gggccacggc 900
taccaccccg aggggtttgt gtgtagccag tgtgggaagg tccctggaaga gggtaggctc 960
tttgaggaga agggcgccat cttctgcaca ccatgctatg acgtgcgcta tgcacccagc 1020
tgtgcgaagt gcaggaagaa gattacaggg gagatcatgc acgcccagaa gatgacctg 1080
cagctgcact gctttacctg tgcctgctgc aagacggcca tccgggaacg ggccttctac 1140
atggaggagg gcgtgcacct ttgcagagga gactatgaga agatgtttgg caggaatgac 1200
catggctgtg acttcaagat cgaagctggg gaccgcttcc tggaggccct gggcttcagc 1260
tggcatgaca cctgcttctg ctgtgcgata tgtcagatca acctggaagg aaagaccttc 1320
tactccaaag aggcagggcc tctctgcaag agccatgctt tctctcatgt gtgagccct 1380
tctgcccaca gctgcgcggg tggcccttag cctgaggggg ctggagtgtt ggccttgcac 1440
ttctgggttag ggctggcaat ggttgacctt accctggctc ctggcccgag cctgggctcc 1500
cgggccccctg cccacccacc ttatctctcc accccactcc ctacaccacc acagcacacc 1560
ggtgctggcc acaccagccc cctttcacct ccagtgcacc aataaacctg taccagctg 1620

```

&lt;210&gt; 33

&lt;211&gt; 1665

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 33

```

cgacgcagag caggccctg gccgggcca gacggagccg gcatcatgga ttcttcaag 60
gtagtgttgg aggggcccgc accttggggc ttccggctgc aagggggcaa ggaactcaat 120
gtgcccctct ccatttcccg gtcactcct gggggcacaag cggcgaggcc cggagtggcc 180
gtgggtgact ggggtgtgag catcgatggc gagaatggcg gtgacctcac acacatcgaa 240
gctcagaaac agatccgggc ctgcccggag cgcctcagcc tgggcctcag cagggccag 300
ccggttcaga gcaaacccga gaaggcctcc gcccgcggcg cggacccctc ggggtacacc 360
tttgacccca gcgtctccct caacaagacg gcccgccctt ttggggcgcc ccgcccgcct 420
gacagcgccc cgcaacagaa tggacagcgg ctccgacggc tgggtccaga tgcagcag 480
caggcgctga tggagaacac agaggactgg cggccggggc cggggacagg ccagtccgct 540
tccctccgca tccctgccc cctcagaggg acgaggtca tgcagacccc ggatgaggag 600
cccttgaaag aatcaagcca ggtgcccagg acagaagccc cagccccagc ctcatctaga 660
cccaggagc cctggcctgg cctacccgcc ccagaccta ccagccgccc gccctgagct 720
gtggaccttg cgtttgcaga gcgtatgccc ccggacaaa cagacacagt gctgacccgg 780
cacagccagc cggccacgac cagcccgctg cagagccgca cctccattgt gcaggcagct 840
gcccgggggg tgcaggagg gggcagcaac aacggcaaga ctcccggtg tccaccgtgc 900
cacagggtca tccggggcgg ctacctgggt gogttgggccc acggtacca cccggaggag 960
tttgtgtgta gccagtgtgg gaaggtccct gaagagggtg gcttctttga ggagaggggc 1020
gccatcttct gccaccatg ctatgaagtg cgtatgcac ccagtgtgct caagtgcag 1080
aagaagatta caggcgagat catgcacgac ctgaagatga cctggcagct gcaactgctt 1140
acctgtgctg cctgcaagac gcccatccgg aacagggctt tctacatgga ggaaggcggt 1200

```

```

ccctattgcg agcgagacta tgagaagatg tttggcacga aatgccatgg ctgtgacttc 1260
aagatcgacg ctggggacgg ctctctggag gccctgggct tcaagtggca tgacacctgc 1320
ttcgtctgtg cgatacgtca gatcaacctg gaaggaaaga ccttctactc caagaaggac 1380
aggcctctct gcaagagcca tgccttctct catgtgtgag ccccttctgc ccacagctgc 1440
cgcggtggcc cctagcctga ggggcttggg gtctgtggcc tgcatttctg ggtagggtctg 1500
gcactggttg ccttaacctt ggtcctctgg ccgagcctgg gctcccgggc cctbgcccac 1560
ccaccttacc ctcccacccc actcctctca ccaaccacgc acaccggctg tggccacacc 1620
agcccccttt caactccagt gccacaataa acctgtaccc agctg 1665

```

&lt;210&gt; 34

&lt;211&gt; 223

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 34

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1             5             10             15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
          20             25             30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
          35             40             45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
          50             55             60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
          65             70             75             80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
          85             90             95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
          100            105            110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
          115            120            125

```

```

Pro Gln Gln Asn Gly Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser
          130            135            140

```

```

Lys Gln Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly
          145            150            155            160

```

```

Thr Gly Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr
          165            170            175

```

```

Glu Phe Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln
          180            185            190

```

```

Val Pro Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu
          195            200            205

```

```

Pro Trp Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro

```

210 215 220

<210> 35  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 35  
 gagcgggcac catggattcc 20

<210> 36  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 gctgcctgca caatggaggt 20

<210> 37  
 <211> 1456  
 <212> DNA  
 <213> Homo sapiens

<400> 37  
 cagcgcagag cagcgccttg gccgggccaa gcaggagcgg gcattcatgga ttccttcaag 60  
 gtatgtcttg aggggccagc accttggggc ttcggctgtr aggggggcaa ggacttcaat 120  
 gtgcctctct ccatcttcgg gctcactcct gggggcaaag cggcgagggc cggagtggcc 180  
 gtgggtgaat ggggtctgag catcgatggg gagaatgcgg gtatgcctcac acacatcgaa 240  
 gctcagaaca agatccgggc ctgcggggag cgcctcagac tgggcctcag caggggccag 300  
 ccggttcaga gcaaacggca gaaggtgcag acccctgaca aacagccgct ccgacgctg 360  
 gtcccagatg ccagcaagca ggggtgatg gaggacagag aggactggcg gccgcggcgg 420  
 gggacaggcc agtcggttc ctccgcctc ctgcccacc tccagggcaa cgagtccatg 480  
 caagaccggg atgaggagca cctgaagaaa tcagccaggg tgcctaggac agaagcccca 540  
 gccccagcct catctacacc ccaggagccc tggcctggcc ctacggcccc cagccctacc 600  
 agccgccccg cctgggctgt ggaacctgag ttctgcgagc gctatgcccc ggacaaaacy 660  
 agcacagtgc tgaccgggca cagccagccg gccacgcccc cgcgctgna gagccgcacc 720  
 tccattgtgc aggcagctgc cggaggggtg ccaggagggg gcagcaacaa cggcaagact 780  
 cccgtgtgtc accagtggca caaggtcatc cggggccgct acctggtggc gttggggcac 840  
 ggtaccaccc cggaggagtt tgtgtgtagc cagtgtggga aggtcctgga agagggtgac 900  
 ttctttgagg agaggggcgc catcttctgc ccaccatgct atgacgtgag ctatgcaccc 960  
 agctgtgcca agtgcaagaa gaagattaca ggcagatca tgcacgacct gaagatgacc 1020  
 tggcagctgc actgctttac ctgtgctgcc tgcagacgc ccatcgggaa cagggccttc 1080  
 taactggagg agggcgtgcc ctattggag cggagctatg agaagatgtt tggcagaaa 1140  
 tgccatggct tgacttcaa gatcgacgt ggggaccgct tcttgagggc cctgggcttc 1200  
 agctggcatg acacctgctt cgtctgtgag atatgtcaga tcaacctgga aggaagagcc 1260  
 ttctaactca agagggacag gcctctctgc aagagccatg ccttctctca tgtgtgagcc 1320  
 ccttctgccc acagctgccc cggtagcccc tagcctgagg ggccctggag cgtggccctg 1380  
 catttctggg tgggctggc aatgggtgac ttaacctgag ctectggccc gagcctgggc 1440  
 tccggggccc tgccca 1456

<210> 38  
 <211> 423

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 35

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe  
 1 5 10 15

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg  
 20 25 30

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp  
 35 40 45

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile  
 50 55 60

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly  
 65 70 75 80

Leu Ser Arg Ala Gln Pro Val Gln Asn Lys Pro Gln Lys Val Gln Thr  
 85 90 95

Pro Asp Lys Gln Pro Leu Arg Pro Leu Val Pro Asp Ala Ser Lys Gln  
 100 105 110

Arg Leu Met Glu Asn Thr Glu Asp Trp Arg Pro Arg Pro Gly Thr Gly  
 115 120 125

Gln Ser Arg Ser Phe Arg Ile Leu Ala His Leu Thr Gly Thr Glu Phe  
 130 135 140

Met Gln Asp Pro Asp Glu Glu His Leu Lys Lys Ser Ser Gln Val Pro  
 145 150 155 160

Arg Thr Glu Ala Pro Ala Pro Ala Ser Ser Thr Pro Gln Glu Pro Trp  
 165 170 175

Pro Gly Pro Thr Ala Pro Ser Pro Thr Ser Arg Pro Pro Trp Ala Val  
 180 185 190

Asp Pro Ala Phe Ala Glu Arg Tyr Ala Pro Asp Lys Thr Ser Thr Val  
 195 200 205

Leu Thr Arg His Ser Gln Pro Ala Thr Pro Thr Pro Leu Gln Ser Arg  
 210 215 220

Thr Ser Ile Val Gln Ala Ala Ala Gly Gly Val Pro Gly Gly Gly Ser  
 225 230 235 240

Asn Asn Gly Lys Thr Pro Val Cys His Gln Cys His Gln Val Ile Arg  
 245 250 255

Ala Arg Tyr Leu Val Ala Leu Gly His Ala Tyr His Pro Glu Glu Phe  
 260 265 270

Val Cys Ser Gln Cys Gly Lys Val Leu Glu Glu Gly Gly Phe Phe Glu

275	280	285
Glu Lys Gly Ala Ile Phe Cys Pro Pro Cys Tyr Asp Val Arg Tyr Ala		
290	295	300
Pro Ser Cys Ala Lys Cys Lys Lys Lys Ile Thr Gly Glu Ile Met His		
305	310	315 320
Ala Leu Lys Met Thr Trp His Val Leu Cys Phe Thr Cys Ala Ala Cys		
	325	330 335
Lys Thr Pro Ile Arg Asn Arg Ala Phe Tyr Met Glu Glu Gly Val Pro		
	340	345 350
Tyr Cys Glu Arg Asp Tyr Glu Lys Met Phe Gly Thr Lys Cys Gln Trp		
	355	360 365
Cys Asp Phe Lys Ile Asp Ala Gly Asp Arg Phe Leu Glu Ala Leu Gly		
	370	375 380
Phe Ser Trp His Asp Thr Cys Phe Val Cys Ala Ile Cys Gln Ile Asn		
	385	390 395 400
Leu Glu Gly Lys Thr Phe Tyr Ser Lys Lys Asp Arg Pro Leu Cys Lys		
	405	410 415
Ser His Ala Phe Ser His Val		
	420	

&lt;210&gt; 39

&lt;211&gt; 1575

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 39

```

cgacgcagag cagcgcctg gccgggcaa gcaggagcog gcacatgga ttccctcag 60
gtagtctgg aggggccagc accttggggc ttccggtgc aggggggcaa ggacttcaat 120
gtgcccctct ccatttcccg gctcactcct gggggcaaa cggcgccagg cggagtggcc 180
gtgggtgact ggggtctgag catcgatggc gagaatgagg gtagcctcac acacatcgaa 240
gctcagaaca agatccgggc ctgcggggag ccctcagcc tgggcctcag caggggccag 300
ccggttcaga gcaaacggca gaaggcctcc gccccggcgg cggaccctcc gcggtacacc 360
tttgacccc gctctccct caacaagacg gccgggccc ttggggcgcc cccgcctgct 420
gacagcgccc cgcaacagaa tgggtgcaga cccctgacaa acagccgctc cgaccgctgg 480
tcccagatgc cagcaagcag cggctgatgg agaaccagaa ggactggcgg ccggggccgg 540
ggacaggcca gtgcggttcc ttccgcatcc ttgccacct cacaggcacc gacttcatgc 600
aagaccggga tgaggagcac ctgaagaaat caagccagg gcccaggaca gaagccccag 660
ccccagctc atctacccc caggagccct ggcctggccc taccgcccc agccctacca 720
gccgcccgc ctgggctgtg gacctgcgt ttgccgagc ctatgccccc gacaaaacga 780
gcacagtgt gaccgggac agccagccgg ccacgcccac gccgctgcag agccgcacct 840
ccattgtgca ggcagctgcc ggagggtgc caggaggggg cagcaacaac ggcaagactc 900
ccgtgtgtca ccagtgcac aaggatcatc ggggcccgtc cctggtggcg ttgggcccag 960
cgtaccaccc ggaggagttt gtgtgtagcc agtgtgggaa ggtcctggaa gagggtggct 1020
cttttgagg gaagggcgcc ctcttctgcc caccatgcta tgagtgagg tatgaccca 1080
gctgtgcaa gtgcaagaag aagattacag gcgagatcat gcaogccctg aagatgacct 1140
ggcactgca ctgcttacc tgtgtgccc gcaagacgc catccggaa agggccttct 1200

```

```

acatggagga gggcgtgcc tattgcgagc gagactatga gaagatgttt ggcacgaaat 1260
gccatggctg tgacttcaag atcgacgctg gggaccgctt cctggaggcc ctgggcttca 1320
gctggcatga cactgcttc gtctgtgca tatgtcagat caacctggaa ggaaagacct 1380
tctactccaa gaaggacagg cctctctgca agagccatgc ctctctcat gtgtgagccc 1440
cttctgccc cagctgcgc ggtggccct agcctgaggg gcctggagtc gtggccctgc 1500
attctgggt agggtggca atggttgcct tascctggc tctggccc agcctgggt 1560
cccgggccct gccc 1575

```

&lt;210&gt; 40

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

```

Met Asp Ser Phe Lys Val Val Leu Glu Gly Pro Ala Pro Trp Gly Phe
  1              5              10              15

```

```

Arg Leu Gln Gly Gly Lys Asp Phe Asn Val Pro Leu Ser Ile Ser Arg
      20              25              30

```

```

Leu Thr Pro Gly Gly Lys Ala Ala Gln Ala Gly Val Ala Val Gly Asp
      35              40              45

```

```

Trp Val Leu Ser Ile Asp Gly Glu Asn Ala Gly Ser Leu Thr His Ile
      50              55              60

```

```

Glu Ala Gln Asn Lys Ile Arg Ala Cys Gly Glu Arg Leu Ser Leu Gly
      65              70              75              80

```

```

Leu Ser Arg Ala Gln Pro Val Gln Ser Lys Pro Gln Lys Ala Ser Ala
      85              90              95

```

```

Pro Ala Ala Asp Pro Pro Arg Tyr Thr Phe Ala Pro Ser Val Ser Leu
      100              105              110

```

```

Asn Lys Thr Ala Arg Pro Phe Gly Ala Pro Pro Pro Ala Asp Ser Ala
      115              120              125

```

```

Pro Gln Gln Asn Gly Cys Arg Pro Leu Thr Asn Ser Arg Ser Asp Arg
      130              135              140

```

```

Trp Ser Gln Met Pro Ala Ser Ser Gly
      145              150

```

&lt;210&gt; 41

&lt;211&gt; 24740

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 41

```

nnnnmtgta ttttatcata ttttaaaat caaaaaacaa aaggcagttg aggttaggca 60
tggaggttgg tggctgtaat cccagacact tgggaagccg aagcacgtgg atcacctgag 120
gtcaggagtt ccagaccagc ctgcacaata tggtaaaacc ctgtctctac taaaaataca 180
aaaaattagc caggcatggg ggtgggcacc tgtaattcca gctacttggg agactgagge 240

```

agggagaatca cttanaaccocy gggggcggggc tggggcggggc gggcctatgcc tggtaatccca 300  
 gcacttttggg agggccgagac agggcgatca tgagggtcagg agatccgagat catcctgggt 360  
 aacatgggtga aacccccatct ctactaaaaa tacaaaaaaa attagccagg cctgggtggcg 420  
 gggacctgtga gtccccagcta cttggggggc tgaggccagg gaaatggcgtg aacctggggag 480  
 gggcggttgc agtgagccca gatcgccca ctgcactcca gcccggggcga caagagtggag 540  
 actccatctt aaagaaaaaa aacaaaccgg gggggcggga attgcagtca gccgagatct 600  
 cggcattgca ctcaagtatg ggtgacagag caagactcca tgtcaaaaaa aaaggcagtt 660  
 gacaggagca agggagcctgg tgaggaaagct gtggcatttg acccggtgtg gttgctatgg 720  
 gccagggttg tgcctagtga gggagctgag gggaaagagc acagggggaca tgcctgaaggc 780  
 ctgggtgttg ggtatgaggca gagatcgggg gcaccttgc gggctcatagc aggtggctgt 840  
 ggtgagatgg aggaagacac ctgggggtact gctctaggct gtcagacata cagaagctgg 900  
 cccagccaaag cccaggggct gcaaggagca tccctttgtg tccccagtg tctgcagctc 960  
 tcagacaccc tcaagccacg tgcctcttgc ccagcccagc actctcagtg gggagccagg 1020  
 tgggagaaaca ggcctcggaag gggacctaag ctatagcagc gagccgggca aggtcggaac 1080  
 tggagccca gcccctggat gccccttggc ttgttgggtt ctgggatact gaggggaggg 1140  
 gacagggtat gggagtgcgg tgcctctacc ttgtacttga actcattccc caggggacag 1200  
 gggagggtct ctccaggtac acagatgccc agtctccca gaggggcctg gtccccatgg 1260  
 aggaabaatc catctactcc tccctggcagg aaggtaaagt gggagacgtg caaggggcag 1320  
 ctccagccccc caacccaggg gctgggtctt ttctggactg accggagctg cctggccacc 1380  
 tgcacacag ggcaggttct ccgtggtgtg tccagaggact gaggccgcca cccgtgccc 1440  
 gctgaagggg ccggccctgc tgggtgctgg cccagagccc atccagctga gggaggccaa 1500  
 ggcacccagg cccctctacc ctggccctac caactcctgc gcaagttcgg ctccgacaa 1560  
 gtgaggtgca ggggtggga aggttgaagg gctgacagcc tggaccctcc tggtaatccc 1620  
 caccogtgtg cctgtgtccc agggcgtgtt cccctttgag gcgggcctg cctgacaccc 1680  
 ggggtgagggc ctctttgccc tccagacccc ctgtgcccct gacctgtgca gggctgtggc 1740  
 cggggccatc gcccgcagcg gggcggctg ccagagctga ccaggcccca gccctgcccc 1800  
 ctgcccaggg ccacccctct gcccctcctg gacacccccc gagagctctg gggatgcca 1860  
 ccaggacctg agccacccac gtcccaggaa atgacacctg ccagagcccg accccagagc 1920  
 ctgcccctac tgcctaggccc gggagccaac gatctgggt ccgggctcta cgttccagtg 1980  
 tgcagagctg ccagtgggccc ccaggccaat gagcacctct atgagaacct gtgtgtgtg 2040  
 gaggccagcc ccacgctgca cgggtgggga cctgagccgc aaggggccc cggagccgc 2100  
 agcccccaca ccagtcccat ctaccacac cggccaggact tgagctggcc cggccgggccc 2160  
 aacgacagta cccctggagg ccagtaacgg cggctgctgg agctggatca ggtggagggg 2220  
 acaggccccc ctgaccccca ggcaggttcc aaggccaaag tgggtgacct gctgagctgt 2280  
 gagcgaggga agggccccc cctctgtgac cggccctgaa cggccagcag agtgggtggc 2340  
 agagggggga ggtgctcccc ctgggacagg aggttgggct ggtgggcaaa cattgggccc 2400  
 atgcagacac agcctgtgtt ccaacctggc ctgaggaac aaggcaggcc gctgtgtgg 2460  
 gacctcagcc ctgcccctgc cccctcatga atagtgtgca gactcacaga taataaagct 2520  
 cagagcagct cccggcagg ggcactcaagg cacaagcccc tgcccacgtt cattgcccgc 2580  
 aacacaagca cccctgtgccc gttccagggg cagaggtgac ctgggccccta cctgcccacc 2640  
 gtgggtctca acccaactga gcagacagac gggatggaaa tcaattaggac tccatgttgc 2700  
 tctgacaggg cagctgacac gaagaggcag ggcgaggagg ctgtgaggct tacttgtcag 2760  
 actcaggagg gaggcaacat agggcccaac tggagacccg gaggcccgag ctgggaggag 2820  
 ggcgtggggg cgggtgtgag gtggagggga ttccagagac accctcgtcc aaaaacattc 2880  
 ttccctgtg aaactccaac aatttgcaga tacttctggg aaccccaggc gtcagtctcc 2940  
 tcatctgtaa aggagagaga accgagtacg tatcaggcat aatccttgat gagggttgc 3000  
 tgggtgcta ctccagtgcca ggcctggggg agacagccg cctgaagact tccatgttgc 3060  
 gttccctct ctgctcctgc tccctggccc ctccagggag cccagccatgc acccaagcgt 3120  
 catgcccaca gggctctgac tggccagccc ctccagcaca cccacccca gctgcacaca 3180  
 cccctggcac atcctgaacc tcaatttctc gacggacaca caatttttgc tctctcctgt 3240  
 ccaagcctca tccctbggcc gcccctcct tccagctcac ttcccttagt ggggcagta 3300  
 cggccctgc ctaggcatgt ccagctgccc ggaaccttt ctggctcttc gaggcctctg 3360  
 cccaccatcc cctctttgtt ctccatagtc ccttccccc gttctctctc gtttcatctt 3420  
 actggtctgg caaagtcacc ggccttgggc gagccagacc tccctagctg ctgcacacag 3480  
 ctgcccacag ccaggagaa ccaattaaag agactgccc catccttctt aacagtgcac 3540

```

ggcagggaact cccggaacaa agagaccctg tcccctagta gggcagcttt tctcctcccc 3660
agaacotcct gtctatcccc acccaatgtc toctcacagg catattgggg aaacagggtca 3720
ggctctccca ccgtatctgc aagtgtactg gcacccatct gtcttcttcc taccctaca 3780
gtagaacacag tgtctgtccc cagctgtgct ctgacccogg ctcttttcc ctcagagctt 3840
ggaaaattga gctgtcccca ctctctctct gcccattca tctaccagc agcttttcca 3900
gccacacgca aacatgtctc gtaatttcc attttaaac ttccttgac ctacatctcc 3960
tcttoggcca cctctgttcc tctgtctctc ttccagcaa aaactgttca aaagagtgt 4020
tgattacttt catttccact ttctccccc cattctctcc tcaattaaact ctcttctac 4080
cccattgtgc cattatgtgg cttttattag agtcacaaac cttattctcc aaacaaaaag 4140
caacaaggac ttgtacttct cagcagcact cagctctggg tcttgaacaa ccccggttac 4200
ttgtatttcc tctacctcca caacaatctc ctccccagcc tctactgctg ccttctctga 4260
gttcttccca ggttcttagg ctcatgtgta gtgtagctca accctgttac acaaaagaa 4320
tcttgaaagc ctgtaaaaat gtccatgcat gttctgtgag tgatctacca agaaaataaa 4380
aaatttttaa aatcaaatgc ccatgctggg gccacacgca aggggtctg atttcatcag 4440
tctggttagt ggggtctggg catccacgct cactggattt ccggtatgatt gtagtatgca 4500
gootaggctg ggaaccactg gcttcagcaa gccagtcatt ctccagggtg cacagaccc 4560
ctagggtgcta atgaccccca aggtctgtct tcaagtgcaca cctcccccctg agctccagat 4620
ttaggaatcc cactgcacac gagacatctg gatgtggaaa agacatctcc agatcccat 4680
ggtgaaaggg ggttggggga atggagactc gtgttcttcc aggtatgtgt tggacacaga 4740
atcnaagcct tggagggatg cttagacat agggagggag atttggctc ctctattcat 4800
gcagcactc cctgatgggt aaggctctag agcaagctga ggccaagagg cgggagctc 4860
aggtgtctgt gccggcaacc cactcccta cagtggcaag cccaagcca gcccttggca 4920
gtccaaatcc caggacacgc tgaaggtcac ccagagagtc aggggtatgg ctagaacccag 4980
aaaccaggac tctggggacc cagcctggca tcttctctct cattacaaat ctgagctgct 5040
ttgtttctca gggatttctg tgatattcca aggggactgt gggaaagaaa gtcttggaa 5100
accaccagga cgttagaggc ctggcctgga gccacaggag tctogggcac cagaggggc 5160
tgggtctctg tccaggtcca gttgctacgc aagggtctgc tgtgtggga ggctcccccag 5220
gggacacaga ccagagcctc gccaccagcc aaggaatggg agcctggggg gtctatttgc 5280
ggaggaactg caggaccccc aggtctgcgc ctcttctct gtctatttgc tgttctactt 5340
tgtcaatcct tcttctctc gtgtgtctat tccatctcac tgbtgbtgg ccttggggaa 5400
atgttagata agacacatta gctgtgtgtc ttcatgttcc taacaaagaa cacaccctgg 5460
aaagagcacc gcagagagtc cccattcccc catctccctc caccatgga atctggagat 5520
gccttttcca catccagatg tctctgtgtc tgtgggattc ttaataaaac aaacatttca 5580
tacagaatgt gagatgatg agatgctatg gggaaaagta aagcagaggg agggcctagt 5640
gtgtgatgcg ggtgagggat ccagggatg ctgtttcagc tgtgatcagg aaaggccctg 5700
ggaggaggcc acatctgagc agagacctaa ataaagtgtg aaactgttgc ctgagatatc 5760
tggagaagtg tttcaagggc cgggcaaccg gcctgtgtgc tccgcctgtc aatcccagca 5820
ctttgggagg ccagggcagg tggatcgtcg gaggtcagga gtttgagagc agctgacca 5880
acatggagaa accccatctc tactaaactc ataaaaatta tccgggcatg gtggttcatg 5940
cctgtagtcc cagctactcg gtaggttgag gcaggagaat cacttgaacc tgggaggcag 6000
aggttgcagc aagccgagat caccaccctg cactccagcc tggatgacag agcagactc 6060
cgtctcaaaa aaaaaaaaga aaagaaaaaa gaaaaaaaaa gaaaagtgtt tcaagcaggg 6120
gaactggcaa gtggagaggc cctgaggcag aaatatgctt ggcctgtctg aggaatgtg 6180
agtggagggg tcagggtggc tggagtggag ggagcagtg gtaggagtc gaccagttt 6240
attcatattc tgtaggctct aaggacttca gttttatttt gagtgcata tgagccact 6300
ggaatgctaa aagctgagag tgacatggtg ctgtgattct ggttttaaaa atactactt 6360
ggctgtctcg tgaagactct ggaaggggca aggttgaag caggatgccc cgttaggaga 6420
ccgttaccag ggcgagggca caaatggca gtggtggga caatggtggc agcagcgggt 6480
agatgtgaac atgttgaagg tggaaatttg agaactctgg ggaggacaga agagaaagg 6540
taacttctac gtttctgctg aaccagtgtg ataatgttg gtggcacttc ttgaagtga 6600
gaaggagtta ggaagggtgg aaaggcaca gtttgaattg ggccatgatg gtctgagata 6660
cctagtacag tggttcccca acctttttgg cagaagggac cgcttctatg gaagacaatt 6720
tttccacaga ctgggggtgg ggtggggatg gtttcaagg gtctcagag 6780
atcattagac tctttttttt ttttttttt tgagatagag tctcgctctg tcaccacac 6840
tggagtgcag tggagccatc ttggctcact acnaccctc ctgcccagggt tcaagtcatt 6900
ctcctgcctc agcctctcaa gtgctggga ttataggcat atgcgccacc accccagct 6960

```



aattttttgtat ttttttagtag agacgggggtt tcacratatt ggccaggatg gtctcgaact 7020  
oatgacctca agtgatccctc ccccgccctca acctcccaaa gtgctgggggt tacaggcgtg 7080  
aaccactgca cccggcccat ttatcattag attctcataa ggaatgagca acctagatcc 7140  
ctcgcatgca cagttcacaa tagggttcac gctcctatgg gagtctaatg ctgcccgtgc 7200  
actcagcttc tctggcttgc cgtctgtcac ctctgtctgt gcagcccatg tctaacagg 7260  
ccacaaacgg ggagttgggg acccctgac tagtaaacat ctaggcaggg ttttggataa 7320  
tgaggttaga gttcctgggg agaggtcagg ctggccatga aacatgggat gcccttgcct 7380  
ataggtgggtg ttgaaagcca caggacagta cggggtctca gggggtgagc ataaagagag 7440  
gcyacatcag atggccaagg ccagagggcag aggaggtatg gaaggagggg ccagtgggggc 7500  
aggggggaagc tgtgaagcca gggaaaaagg gtgttctggg gaaaaggatc aacctggacc 7560  
agtgtctgcc ctaggcaggg caggatgaaa cttaaccacc accgattcca cggcccatg 7620  
gctccaggc cacaggggac cttgagaaga gagatctcag gggacgggtg cggacaagag 7680  
ccgcctggc atggcttcaa gagatcaact aggaagaaagc agtgagagc cgataaacag 7740  
acaactccct ggaggaattt tactctcgag agyagaatta agggtagta gctggagagg 7800  
gatgtggggc caagagaagg tcttttaacga cagaaactct caggcgggtt tgtgcagaac 7860  
agggctgggtg tgatgactgt ggtatggagag gggagaactg cagcagactc gtctaggag 7920  
gaggtgatgg gcggggacca ccaagcaggt cgaagggtgga cgcctcttcc ctaccccca 7980  
caccgcagtg tgcctcagtg cgtgcggcc ggccctagtg cctgggctga acgaggggcc 8040  
gggactctga ggaagccctc caggcgccga gtccgtctgg ccaaggtgga gcgggacggc 8100  
ngcttccgac ggtgcggggg tgggtcggg gttgcaggga catccggcgt ccgctcctgc 8160  
cctgttttcc tgccttcgca gagcgttggg caactctagc tttaaacgac cctgtccccc 8220  
tcaacttgte tccccagcc cctctgattt acagattctg cagtccccga gggttgcggc 8280  
tacgataccg aactctcggg cagcctgcga ggcaggtatg atcgtcccat ttttggagt 8340  
agcaaacctaa ggttcagaga ctactatgtc ccaggtcggc ctggtttgaa ggtccgctct 8400  
cctctccctc cgcagcggg cgttcgagag gactgggcga ggcagcgtc cctaaaggag 8460  
gcgacccgca gccccggccc cctcccgact cgcctccgtt gcaggggccg ggtcggcgag 8520  
gctctcagc tctaaagccg aggggacttg gtgattgggc aggacgggaag agctgggttg 8580  
ggctttccac cagcggagaa agtctagtg gctgtggtgc gacgagggcg tggcctggcg 8640  
cccgccccc gtcgagggc tcaaaagtga ggttggtctg gggggcgggg tcagaacact 8700  
ggcggccgat ccaacgagg ctccctggag ccgacgcag agcagcggcc tggcggggcc 8760  
agcagggctat cgaagacgc gggggtgga ttgggctgga ccaggcgggc gcccggggcc 8820  
tgctgaggac cacaaagggc actgggggtc gtggtccagg ctgtgcttcc tcccgctggc 8880  
ctgggcccct gctccggcc cggcccccgc ctctctgccc cttaagccggc tggggcgggg 8940  
cagattggcg cctgcccgt tccctgcccg gggccagctc aatgcctggg gcccgggcgg 9000  
gggactaagg ggaactgag tcaagtctgg gtgggagcag ttctgtgtgg gaggcaccac 9060  
ccccactgg gctcggggaa ggtatccctc ccaagctatg cttaggggtc ccagccccc 9120  
ttgtctcca caggggcggc accccactcc cgccttccc ttcttcagca ccaggggtc 9180  
cggccctggc tccagcagc ctgcactggc ccgggaatgg ctaggaggat ccgctgcagc 9240  
cgcctccctc cctccctc cctccctc cctccctc cctccctc cctccctc 9300  
ccccctggc cccaaagccc cgtgtgctcc ctccgtggc tctccgcaca gtgtcagctt 9360  
acacgcctta tatagtcga gcaggctcca gccgagggct gctgcccggg cctggggggg 9420  
ggggagagga gaggcggccc ctgactcacc cggacccccc gaggtccag gctggccttg 9480  
ggggagggcg cgcagttta gtcctcggc ccaacccctg ttgcaagaa cctcaagcct 9540  
ggattcaggg acccctcacc gttccagtc caaggggagg ggggctgtc ctgtcttcc 9600  
aaagttaggt ccgacggcca gcagccagc ccagcctgac aaataacctg cctccatgg 9660  
cttgggggtg ctcaagggt gcccgtgct gctggccc accatgctgg cactcactt tctcttct 9720  
agctggcccg gctgcctgc ctgcccgc caaatgctg caaagttagt ctggaggggc cagcacttg 9780  
cctctcagga gccggcatca tggattcctt caatgtgccc ctctccattt ccgggggtgag 9840  
gggcttccgg ctgcaagggg gcaaggactt gctcttctgg gcttaggtct ccagagggtg 9900  
cctaggtttg gggagggggc tccccagcg gggccagcag gcraggagat gtctaggtcc 10020  
atggggggag tctaacagg agctggtcag gactcccaat cccgatgctg gagcctaaca tttcaggagg 10080  
ggagatgtag tggtaacctg ctgcacaaag cccgatgctg gaggctaca tttcaggagg 10140  
gcaaccagge ttctccagc aggtggagg atgcttggag ccaggcccca ggcaggacac tgtgactata 10200  
cacctctgag ggtggaggg atgcttggag ccaggcccca ggcaggacac tgtgactata 10260  
tatctaaact tgaggtggca tgcggccag ggaatagcaa tctcatccc gtttttccag 10320

atgagagaag	aaatcatgga	gaggtggcgg	ggctogctca	togagtcacat	ggtagagcag	10380
ggattggaat	tgaggcacag	catggcgtac	atctttttgtg	ggtagaaggg	gtctctcccc	10440
agcctatgta	aggacccaca	tccactgttc	ccattccagga	tgtaggtggc	tttgaooooa	10500
agcagaagtg	taggacaggg	ctccattcta	ggggcttaac	ttcagcttcc	aaagagcctgc	10560
cctggtgtgg	gtggagctgg	aggctggctc	ctccctgtag	cagggggatt	gccttataag	10620
cccaagaatg	cagccccacy	ctgggtgggc	caacagtggc	tggggtctgc	agagctgaaa	10680
agggtctggc	taggcctggc	ccctgaacc	ccactgggtg	gcctctcagc	tggtcaccag	10740
gctgcagctc	cagctgtatg	gtccagtgtg	gagacacaa	aaattgcctg	ccagagtggt	10800
gtgaggccag	cctgtcggct	ggcatctctg	actggcctgg	gggtcaggag	ggggtgggga	10860
cttctctccc	ctatctccgc	ctgccccggg	agacccaccc	aggcggcggg	tggtcaggca	10920
gctgttgtca	ggaagcccaa	ggcaagccca	gcctggaggg	gcccagaggg	togtggcctg	10980
aggaggggct	caagctggag	tctgtctgta	ggagctgggc	gtgggggtta	gggtgggcag	11040
gccagcagtg	ctctctcag	gggtcctttg	atggcattct	cctgggaacct	gcccogccag	11100
cagggtagtg	aggcagtggt	tgccctatga	cacaogtccc	actacatagc	cctcacacag	11160
ccctgaaccc	tacctgacgt	cctgctccct	gggaaggtgc	tggtccagtg	tgtctgggga	11220
gcctgaacct	cagtttcttc	cctgabtggag	atgactttca	gatattggct	gttgggggca	11280
ctccgggctc	cagctccctg	gtcagcatcc	ctggcatgtg	ggcgggggca	ctagctgctc	11340
ccagccctgg	agttggacct	gggcccacat	gggtgggtga	ggtagggctt	tctgagttag	11400
gccagccccc	tcccccctcc	ctgaccccag	aattggagga	ggtagggagg	gcaagggctg	11460
gctgtggtcc	caggcctggg	agatgaggtg	acgtctggga	ctgggggggt	gggctggctca	11520
ggctgactca	ccccacctc	atgcagggtc	cagccccctg	gctttttccc	tacttgggtc	11580
ctctggcctt	accctgcccc	tggcttgagc	ccctccctgc	ctctctccag	ccacccgccc	11640
agcgtgtgtc	tctgtctctc	tgtgcccctc	cccacgtctc	gaacacccct	cctcctctgt	11700
gcttctgtcc	ctcctcactc	tgggaaggga	agcgtccccc	gccccccacc	ccctctccag	11760
gagccagcta	gctgcacccc	aagaccccca	cctcgggttc	agccccacag	tcccaggagc	11820
cagccctgtg	ggcagggagt	ggctgggcca	ggtttccctt	ctactgactc	accttgacct	11880
tgagtaagtc	acttccctc	tgggtgtgta	cttccccata	cacagtataa	ggggttgatt	11940
tagttggatt	gaactaaagg	tgaggagtg	gctcagggtg	tctccaggtg	ggctgacccc	12000
tcagttgggc	ccccatgtct	agcagaggtg	gcccacagtg	gtggagcctt	agggtcagag	12060
acacttctct	gctctgctct	ttactagctg	ggtagcttga	ggcaagttgt	ttaacctctc	12120
tgtgtacatt	tgcaagtga	aatgggtga	aatccagat	tactccacaa	ggttgtttga	12180
agattcagtg	tcaatatgta	gcatagttgg	tgctcaataa	aetgaagcaa	gtcttcttat	12240
ttagcagtg	aggagggggc	cgcagagctc	tcttagcctt	ctgacctcct	acgcaagcaa	12300
gaggtcattg	tgagccagc	tgcctttctc	tttcccagtg	ctgtcaagct	ctgtgctctg	12360
ctgcccctgc	ctctgacatc	tctctgaaac	ctcttgcttc	ccctctccct	gcctcagctc	12420
agtctgtgca	ctgaacccac	tgaggagcct	cctggggcca	ctggcagcct	ggaccccccc	12480
agatccccc	caccagtgta	aatgtcttcc	cagcactgcc	tcacaaaagc	ctacttgatg	12540
cagtgccagg	cctcttgcca	gatggctggg	tggctcccta	ggcttgagcc	cagtcasgct	12600
gccctgcttg	tgttgctggg	gctgggctag	aggcctggaa	ggggtttatic	agggtcaccc	12660
tctcagggcc	tgggagatac	ccaatcccag	acattaaaac	tgccagtagc	ccctctacct	12720
tcaaaagccaa	gtcctggtcc	cttcccctgg	cattcaaaagc	catagttaagt	gaactctcac	12780
ccgtataggca	gcacacgcca	ttctccttta	ccgaggccca	ccgtctcttc	aaagtcattc	12840
ctgatggtct	cagctcatgc	tgttggaagg	catttctccc	agcctactgt	ctctactcat	12900
tgccacagga	acccgggact	cccagctcaa	gagcctgaag	gattgggggt	aggggaaatt	12960
ggcagtcag	ggcttgggag	tgacagccct	gtatggccta	cgsagtccca	gctgtcaact	13020
taggtcccat	tcaggcagtg	ttcacaggga	accgggagat	aacagggcct	gttccctggct	13080
ctcaataagg	cccagcagac	ccctatagat	ggcccccgac	aggggtgctgg	ggggtgagag	13140
gtccataaga	gcccccggtg	gtttcgggga	ggaaagctgcc	ccctgcctgg	gccagagggc	13200
atatctggta	ggtaggagtg	cctgggcagg	aggccagcag	gagcctcaaa	aggcaatggc	13260
cctcctgaaa	caettgggct	ttagcctgag	cgtggctgtt	tgtggacatc	atagcaattt	13320
ctggactgtg	ggggagggtg	gtggcgttga	atagataagg	atcgtgactg	gggaagctca	13380
ggtagagccc	acctgaggga	gagggctctg	cagtgaaata	ataagcagtg	tgactgggaa	13440
attgtgaagg	tcagggtgag	gccacacact	cctgggttgc	tttagtgtcc	agcagctggc	13500
tagaactatg	ttgaatgag	agctctctgg	gttctgggag	tgggacagct	ttgggtgggg	13560
cagtggttcc	acagtcagcc	tggcttgggt	ctgcagggtc	cagggcctcg	gtcactttgc	13620
ttctctctcc	acagctcact	cctgggggca	aaggggcgca	ggccggagtg	gcctggggtg	13680

actgggtgct gagcatogac ggcgagaatg cgggttagcct cacacacacac gaagctcaga 13740  
 acaagatccg ggcctgcggg gagcgccctca gcctgggcct cagcaggtat ggggtggac 13800  
 atggatgggt ggcgcgcgcg tggcagtggg gatccctgcg gcccgcccgc ctgtcacgct 13860  
 ttccctctcc tccagggccc agccgggttca gaggcaaacg cagaaggtag gaggctggcc 13920  
 gggacatccg ggcggtgggc ggtgtgggct tggacggcca ggctgtctcg cctccctggc 13980  
 acattctcgg taccaccaac cctggccggg agtgaggggc agaaaccgga gctaaggcgg 14040  
 gtctaggggc ctggagttga gccaggggct gctgcacggc cctggcaccg cgcattgtccg 14100  
 cctgtctgtc cgcctgtctg tccgctgct gcctccgcgc gcggcgctg cgtgtctgcc 14160  
 cgcactcggc cagccctcgg tccgtcgtgg actgagatcg ccaactccca atgggocct 14220  
 tgaacctga gtctcctct ccccgtagcc tccasataga tctagggggt ggggtggggg 14280  
 tggggggctg gagctgcgcg tgtcctctgc tgcaggcgcc ccacttccac ccaggccccc 14340  
 accttacctc gccgcgcgcg cctgcgcggc cgtgtctctg cccaggcctc cgcgcccgcc 14400  
 gcggacctc cgcggtacac ctttgcaccc agcgtctccc tcaacaagac ggcccgccct 14460  
 ttgggcgcgc ccgcgcgctg acagcgcccc gcagcagaat gggtaogtcg gccctgcgc 14520  
 gcccgcgccc acgccaatcag gcccaactgt gccccaagcc cgtgcgcgcg tctgtctcag 14580  
 tctgtgtctg gccccagccc ggggggaacg tgcggcaogc cccctggcgg ccgggtggg 14640  
 gctgcaggca caggccccc cccagggtct tggcgccttg cagggcacog cctggggagg 14700  
 ggtctctgaa tgcgcgcgcg cccctgtctg ggcgttgggg gttgggttgt ggtgtcgggc 14760  
 cagctgagcc ccagacactc agtgcgcgct tgtcccgccg tgttctgacc cctcccgctc 14820  
 ttctctctct tccctgtctc gtccctttgt cccctttatc gtctgtctgt cttatttctc 14880  
 tcacagttgc agacccctga caagtacgtg agcccccctc tgcctgtgac ttctctctc 14940  
 cttttggcac tctgggtggc ggcctccccc caccctggct gccctcctct ccaactcgcc 15000  
 ctccctgtct ctacactacc agccagcag ggtccctggc ctacccctta ccaactccct 15060  
 cccactcact taaccacaa cccacatgac caaatcctgg gaggggctgc cccacccg 15120  
 cccccccagt gtgggggtct gagccacacc ctccccacag acagccgctc cgcagcctg 15180  
 tccagatgc cagcaagcag cggctgatgg agaacacaga ggactggcgg ccggggccgg 15240  
 ggaagggcca gtcgcgttcc tccgcctcc ttgcccacct caccggcacc gaggtcagta 15300  
 agtgccagcc cagggcaggg ggtactttcc tgcgccccag cccaggcgtg atccctgac 15360  
 ctgtgtcttt tttggtcaat ggcctgcctc gtctctcag tgcgaagacc ggtgaggag 15420  
 cactgaaga aatcaaggta cagggaaggc caccagcccc tctccacct cctgctctc 15480  
 ccattccagc tactgcctg tgtctactcc tgaggctccc agctggggct ctcaattctc 15540  
 ccttccctcc tccctccct ccttccctcc tccctccct ccttccctcc tcttccctc 15600  
 ccttccctcc tcttccct ctttccctcc tccctccct ccttccctcc ccttccctcc 15660  
 cctgctcccc tccatctct ccttccctcc acttctctc ccttctctc ccttctctc 15720  
 gggaaaagta tgtcctggag ctgcagagcc cagcctacac cgcctccgg gactggcacc 15780  
 accagcgctc tgcaccgtg ctcaagctgc agtcgtagcc cggccctctc cagccggctg 15840  
 cctctctgct ccttctctt ctgttccctc tgcaccaggc accccttag tgcctccagc 15900  
 ttctgcctac ctaccccccc ctttctgtgc cctggcctga gctcctgtc ggcctggccc 15960  
 tggcggccca cctgggttca tctgacactg ccttccctct tgcctctgt gtagctgtg 16020  
 ctgcccaggtc tgtgtgtct tgggcctgga ataaacatc tcagccctgc ttgctctg 16080  
 tgtctctctc ctttgtggac ctggttctga tttgggggtg ggggtgtctt cgtggttcg 16140  
 actgtttggg cctggtctc cttgttttca gtcgggaggg gtactctgga aaggggccct 16200  
 gccctgccat cacagatggc ttccctggct gaggggagcc ccaggagctc cctcagaagc 16260  
 gggagccctc cctgtctcc cagctagaga ccgcacacca gctaactgga cattgtcagg 16320  
 agaagctgcc cttcccatcc ctaccccagc gggacctgga atccaaactc gcagtttcca 16380  
 cgcgccagc cctctccct ggggcagga ggaaccaggc tgggggggtg ggcctgtca 16440  
 ggaagctcag ccatgcagg ccttgaatgg cagactcttg agccaggctc ccaggacaga 16500  
 agccccagcc ccagcctcat ctacacccca ggagccctg cctggtgaga gggagtggc 16560  
 tggggcctgg gcaagggtg gcaagctcca ggggcctg ggtggtgggc ttctctcagc 16620  
 tgcctggggc tccacccccc ccttggggg cccctttaga gtaacttcc 16680  
 ccggcaggcc ctaccgcctc cagccctacc agccgcgcgc cctgggtgt ggacccctgc 16740  
 tttgcgcagc gctatgcccc ggaacaaacg agcacagtgc tgaccgggca cagccagccg 16800  
 gccacgccc cgcgctgca gaggcgcacc tccattgtgc aggcagctgc cggagggggt 16860  
 ccaggagggg gcagcaacaa ccggcaagac cccgtgtgtc accagtcca caaggctc 16920  
 cgtgggtgg cctgttctc tccgacctg gcttcccat cctgcagccc agccccacct 16980  
 gtctgcccac ctgtctgccc ccagctgca ctggggggaa taaggattca gttctcagc 17040

```

ggagtaggag tagggacctg ggctgggtcc tcccattctt aatcccacgc taactacccc 17100
agcccaccca caacaaactgc tagcagcctc tgcctgggag aaatagccga agggccaacc 17160
ataggtcgaa gctgcacccc tacctttgct gctctctggg caasgagggg cctgccccct 17220
cccagcgctg ctgcctctcc ctctgctct ctgtctccct ctgctctcag agcatacagg 17280
cctggagcca cccctctgtg gcactgcccc tgggggcca gcagcatcaa acacccccca 17340
gcacagcgct gccggattct agagccttcc taattcgag gectggcctg ctctcatctc 17400
tgtcagotct tttttttttt tttttgaaac agagtctcac tgtgttgccc acgttggggt 17460
gcagtgggag gatctcggtc cactgcaccc tetgctctct gggttcaga gattctcctg 17520
cctcagcctc ctgagtagct gggattacag gcacccgcca ccatgctctg ctaattttgt 17580
attttttagta gagacggggg tttaccatgt tggccagggt ggtctcaca cctcaccctc 17640
aggtgatctc aggcctgctt tggcctccca aagtgtctgg actacaggtg tgagccactg 17700
tgcccagcgg actctatcag ctcttgccag gtagaacagg caggccagca ggacagggca 17760
gctccagggt ttgcccaggg ggggtctcag ttttatgagg ctccagtctg cagcccttcc 17820
tcccgggggt ctccctgctc taaagctgct tctctgtctc ccagcagttc agtgtggggg 17880
actggctctg taagcttcat ggtgcacag gtcaacttcc aagcctgtct tctatcctat 17940
gtggaaaatg gggagaatga actgtccctc ccaaggcctc ctggtgggtg gtcagtcaac 18000
ctgaaggggg ccaagacccc cactctcttg cgtgtgctcc ctctgacccg tctcgctctc 18060
ctgcaggggg cgtacctggg tggcgctggg ccacgcgtac caccggagg agtttgtgtg 18120
tagccagtgct gggaggggtc tggasagagg tggctcttct gaggagaagg gcgcacatt 18180
ctgcccacca tgcctatgag tggctatgct cccagctgtt gccaaagtga agaagaagat 18240
tacaggcggt agtagggctg gctggggggg aggtgtgtcc aagcctgtca gtgggaacga 18300
gggtgtgtgg gaaccccaca gtccagggtc ctcccggagt gaggctccgg gtccctacca 18360
gggtaataaa tgggtgtgtg taotggctc accctgcatt agtcaggatg ctcttaacaa 18420
atgacatgt tctgtctcag aaacccgcca aggtgcacaa gacagggag accaagccag 18480
gagaagccct gggccctcct gactccact ttgggtctct cctgcccctg tgaatgaca 18540
gaacggccaa cttgacacgc tgaagctgtc ctgtctcatg cgtctctctc atttctggat 18600
ccagagccag ggctgcagg agtagccaga gagctctgtg tgggtgatgt catattagt 18660
aggtttacct tgacccagag cagtgggaaa ctcaaaataa tgggtggctta tttctcatct 18720
aaaaacatcc cgggtgggtt ggtctgggac tbatctgggt gaccagggct ccgccttgtt 18780
gcttgactgt tggcagcacc tgettactta ccactcatgg tgcaagatga cacttcagcc 18840
tcgcaccaaa tgcctacctt ccagccagca ggaagtcgga agpagaagaa aggggacaga 18900
gcccctggc gtccatctct agaggatgt gccactgaa cctctgcttt catctgttg 18960
gtcagaaccc agtcacatga ccacacccag tggcaacgga ggctgggaaa tatagtcttt 19020
atthttggga cccatgtgtc cagcaaaact ggggggttcca tcagtccgca agaaccggag 19080
agtggccgat gcagtggctg atgtctgtat ccagcactt tgggaggtcg aggtgggtag 19140
atcacctgag gtccagggtt caagaccagc ctggccaata tgggtgaacc ctgtctctac 19200
taaaaaataa aaaattagct ggggtgtgtg ggcacactgt agtcccagct acttggggag 19260
ctgagggcgg agaactgctt gatcttgaga ggtggagggt gcagttagcc aagatttgtc 19320
cactgccttc cagcctggga gacagcaaaa aaaaaaaaaa aaaaaaaaaa aaaaagggcc 19380
agggcaggtg gctcacacct gtaatcccag cactttggga ggccgagatg ggccgatcac 19440
gaggtcagga gattgagacc atcctggctc acacggtgaa accccatctc tactaaaaat 19500
acaaaaaaat tggccgggca tgggtggagta gtcccagcta ctggggaggc tgaggcagga 19560
gaatggcgct aacctgggag gcagagcttg cagttagccg agatcgcgcc actgcactcc 19620
agcctgggca acagagcgag actcttgtct caaaaagaaa aaagaaaga gaactctgcc 19680
tcccagcctt gggctcctgc cctaccagcc cacacccctg gtagagcctc ctctccccc 19740
agctcaagc ccaagttcct tcactgtgac ctgtctgtct cctctaaaac aggcacacc 19800
agacagtga agagccagc cagacatggg cagaaaacct attctgtga tctactggct 19860
gtgtgagcag gggctagtgt ctctctctgg gctcactga agagaagggt ggcactatgc 19920
tagggccggc acggttgcaa ggtagatgta agatgggtga cagggtgtgt ggagggcaga 19980
aatgcacct ccgaaggcta catgtcccc acacttatgt cttgcttggc ccacactgtt 20040
tcatttttaa atcagtagca aacaatttaa aaaatcagaa gatttgctg catgatgcag 20100
tggctcatgc ctgtaatccc agcactttgg gaggccaaag tgggaggatt gcttgagccc 20160
aggagttrca gaccagcatg ggcaccatag caagaccctt gtttctacaa aaaaaaaa 20220
attagaaaat tagccaaagt tgggtggcatg cactgtgtgt cccagctact tgggaggcag 20280
agggaaagt agatctcctg ctttttattt ctttatgtat aatgatagg tcttgcctct 20340
tggcccgagg tggagtgag tggcatgat actgtctact gcagcctga tctctggg 20400

```

tcagaggatc	ctcccaacctc	agcctcccaa	atagctagga	ctagagggtgc	ccaccagcat	20460
gctcagcaga	tttttaaatc	tttttgtaga	gatgagggtt	tgtatgttg	ccagagctgg	20520
tctcgaaetc	ctggcctoga	gogatccctc	caccttggec	tcccaaagca	ctgggattac	20580
agacgtgagc	cactggcgc	agcagatttc	tctttaacac	ctagatttca	gcctgagcca	20640
ggcaggcatt	cctgaatgaa	ccagttagta	tgtccacaga	agsagaggtc	ctcctccgtg	20700
tgacacagtc	cccacttgge	ccttgacagg	attggatctg	ggatccctgg	atttaasctc	20760
agggccatcc	tcataacagc	ctcacaaggc	tgggatttag	ttccagtttc	acaagggaag	20820
aaaccaagac	ttgagaaggt	caaggctctg	ccagaccac	acatcttgga	ccctcatacc	20880
gcctcgaggc	cccatgctgc	cctctgcctg	ctccagatgt	gaatactgct	ggccctggct	20940
ggcccccggc	ggcccccagg	gtcctaggga	tgaacagccc	agcccsaggc	gagctcagcc	21000
ccttgctgct	ctgccccttc	ccactcctg	cggaggccag	togactcacc	cacaaagggc	21060
caggcactgt	ggggatagat	cagctaacaa	aacagttgat	gcttccctgc	cttctggggc	21120
ttacattttg	gctggagaa	gaggggagag	gcagactgta	agcaataagg	gcataaagta	21180
gggtgcctgg	aagtaatgtt	agatcacgtt	acggaaaaca	ggaaagagca	gagcgacaag	21240
tgctggggtg	ogtggtgcag	ggaaggcagc	tggctgctgc	tgggtgkgtc	agagtgggcc	21300
ctcatggaga	agactgcatt	cgagcagaaa	cctgaagggg	gtgaggggtg	agcctagaga	21360
tatctggggc	agagcagtc	aggcagaggg	gacagccggt	gtcagcccca	ggacaggagt	21420
gtgctgggtg	tgcaggttcc	aggcagaggg	ccagtgtgca	gaggaagggt	gagaaagcaa	21480
gggagagcag	tggcggagac	gggtgggaac	gaggtcagac	ctgctggcct	ccagcctctg	21540
catggggctt	ggctcttgct	gggagcaatg	ggaagcagta	cacagtttca	tgcaggggga	21600
gaaggccgtg	cttgggttgc	aggggcacgc	tgtggcagct	gggatcagag	agaggagcct	21660
gtaggccagt	tgttatgtgg	tcccacgggc	cagatggcca	tggcttaact	cacttcaggg	21720
aggctgtgag	aagcactcag	actctggatg	tgccttgagg	gtgggcccc	ctggatttcc	21780
tgggtgacct	ggtgtggggk	gtgagaggag	ggtgtgtttg	gctgcagccg	acaggagaaat	21840
ggagttgcca	tccgctgat	ggggatggct	gtgggaggag	aggcttgggg	tgagggaatc	21900
aggaactgag	tgttgacac	ggcaagtctg	aaggcgagct	ggtcgtccac	tcagagacct	21960
tggagttgga	gatggaggtg	tgggagtcct	gacaggttag	atgtagtgtt	taccgcgaga	22020
aggaacaggg	cttgcgccca	gcctcctctg	gttcccgtga	cccagggcag	ggcaggaggg	22080
gcctgagcct	gccgagtgac	tgggaacctc	ttccaggaga	tcatgcaagc	cctgaagatg	22140
acctggcacg	tgcactgctt	tacctgtgct	gcctgcaaga	cgcacatccg	gaacagggcc	22200
ttctacatgg	aggaggcgct	gcctatttgc	gagcgaggtg	cccactggcc	agtgaagggtg	22260
aggagggtatg	gtgcattggg	caggcatgaa	tccaggtcct	ctttctctct	gccccactg	22320
tcagactatg	agaagatggt	tggcagagaa	tgcactggct	gtgaotccca	gatcgacgct	22380
gggagccgct	tccctggggc	cctgggcttc	agctggcatg	acacctgott	cgtctgtgog	22440
gtgagagccc	cgcccccaga	actgagcccc	aagcccaccc	gcctctgttt	cattcccccag	22500
gagatgcagg	aggaagttgg	aaggggcctc	tccgtgtgct	cccaaccccc	tgtgactggg	22560
actttgtctg	ccttagatat	gtcagatcaa	cctggaagg	aagaccttct	actccagaaa	22620
ggacaggcct	ctctgcaaga	gccatgcctt	cttccatgtg	tgagccctct	ctgcccacag	22680
ctgcctgggt	ggcccccagc	ctgaggggct	tggagtctgt	gcctctgact	tctgggtagg	22740
gctggcaatg	gttgccctaa	ccctggctcc	tggcccgagc	ctggggctcc	ctgggcccctg	22800
ccccccccac	cttatcctcc	cacccactcc	cctccaccac	cacagcacac	cgatgctggc	22860
cacacccagc	cccttccacc	tccagtgcac	caataaacct	gtacccagct	gtgtcttgtg	22920
tgccttcccc	ctgtgcaccc	ggaggggcag	aatttgaggc	acgtggcagg	gtggagagta	22980
agatgggttt	cttgggctgg	ccatctgggt	ggtcctcgtg	atgcagacat	ggcgggctca	23040
tgggtagtgg	aggaggtaca	ggcgagaccc	catgtgccag	gcccggtgct	cacagacatg	23100
aggggagcca	ctggtctggc	ctggcttgga	ggttagagaa	gggtagttag	gaagggtagt	23160
tagcatggtg	gtcatgctct	gtgatccag	cactttggaa	ggccaaagggtg	ggcagatcgc	23220
ttgaggtcag	gagttcgaga	cctcatggcc	aacacgggtga	aacagcgtct	ctagtataaa	23280
tacaaaaatt	agccgagtg	ggtggggcat	gcctgtatcc	ccagccactc	aggaggctga	23340
ggcgggaaaa	tcaactgaac	ctgggaagtg	gaggttgacg	tgaactgaga	tcacaccact	23400
gogcgcgagc	ctgggtggca	gatggcagag	cgagacccctg	cttcbaaaaa	aaaaaaaaaa	23460
aaaaaaaaaa	gaagggtagt	tgtagtgggg	ggtggatctg	cagagatctg	gtgtggaaaa	23520
cagcaantggc	cacagcaag	tccctggagg	gcccagctgct	gtccaaacag	aagaaggcag	23580
ggctggagag	ggtagccctt	aggtcctggg	aagccacgag	tgcacggcag	tagagctggg	23640
gctgtctctt	gaggttaggg	cagggcaggg	cacagcagag	tttgaaatag	gttctgttgtg	23700
tattgcagaa	aagaggcccc	agaacactga	gggagtgacg	gagggaggct	gggagggagg	23760

```

gttgacgacg ggccatagggg cgggggcccag gcaagggaggg ggcagagagt aatatggcag 23820
agatgggacc cagtggcagg tccgggggat gagggatgga gagaaggaca ggagcgttgc 23890
caggcatctg ggcataacca gacatgctca cgtgtctcc ccgaacctc ctagcaacct 23940
tgogcogtgy tctgcaatca ctatctcat ttttctttt ttaactttaa ttttttttct 24000
ttttaagaga caggatctcc ctaggttgac cgggctggtt tcaaaactct gggtccaagc 24060
aattcttctt ccttagcccc aaagtgtctg cattacaggt gtgagccacc atgcttgccc 24120
caettatctt ctagatgagg cacagaaaga ttgggagact tgaaccaagg cactgtgtca 24180
ttgagccatg agccagacta gaatccaggc ctgaagctgg gtgogctgtc ccaggactgg 24240
ctggcactga gtaccatttg ccaggcggca tctctctggg aagctgactt ctgcccggta 24300
cctggaggac tctagacctt ggtggtggcg cgtcactct ggggttctt gctccact 24360
gatgcccgca ccacctaga gggactgtca tctctctgt cccaagctg gactggaag 24420
actgaagaga agccttaagt aggcaggac agctcagtgt gccatggctg ccgtctctc 24480
agtggctcct ggcattgagg cctgcaaac acctgttact ctctcaaca ggcccttggc 24540
ccggtccctt ttaagagacg agaagggtcg ggcacggtga ctacacctc taatccagc 24600
actttggaag gctgaggctg gagaagggtt ccagcttagg agttcaggac cagcctgggc 24660
aacatggtga gacctgttt tgttttgtt tttgttttt ttagatggag tctgtctctg 24720
tcgccaggc tggagtgcag 24740

```

&lt;210&gt; 42

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Rattus norvegicus

&lt;400&gt; 42

gcactacctt gaaggaaacc atggc

25